Level Spacing for Non-Monotone Anderson Models

John Z. Imbrie*

Department of Mathematics, University of Virginia

Charlottesville, VA 22904-4137, USA

`imbrie@virginia.edu`

Rajinder Mavi

Department of Mathematics, University of Virginia

Charlottesville, VA 22904-4137, USA

`rsm8y@virginia.edu`

Abstract

We prove localization and probabilistic bounds on the minimum level spacing for a random block Anderson model without monotonicity. Using a sequence of narrowing energy windows and associated Schur complements, we obtain detailed probabilistic information about the microscopic structure of energy levels of the Hamiltonian, as well as the support and decay of eigenfunctions.

Contents

---

1 Introduction

1.1 Background

We present a new method for proving separation of eigenvalues for disordered quantum mechanical Hamiltonians on the lattice, in the strong disorder or weak hopping regime. The traditional approach to this question is based on Minami's method [27]. When available, it provides a bound on the probability of two or more eigenvalues in an interval, and thus gives probabilistic control on the separation of eigenvalues. Variants of the Minami estimate have been an important component of proofs of Poisson statistics for eigenvalues in an interval [27, 29, 22, 19, 4, 9, 17, 16]. They are important for understanding the behavior of the ac conductivity [23]. They also lead to minimum level-spacing estimates and simplicity of the spectrum, provided the probability distribution of the potential is Hölder continuous with index $\alpha > \frac{1}{2}$ [24, 9].

Obtaining Minami-type estimates has been problematic for so-called 'non-monotone' models, in which the variation of the single-site potential with the random variables may change in sign from point to point in space. While there are a number of results on localization for non-monotone models [25, 31, 10, 32, 33, 13, 14, 26, 7, 12], there are only a couple of multi-level results that we are aware of. Simplicity of the spectrum in infinite volume is established in [28]. In [30], a Minami estimate is proven for some non-monotone models that admit a transformation to a monotone situation. In [11] an N-level Minami estimate is proven for a class of random block models, in which a $k \times k$ random block Hamiltonian sits at each site of the lattice, and the blocks are coupled through a deterministic hopping matrix. The randomness must be sufficiently rich to ensure the success of the method. Of particular interest is the case of 2×2 blocks of the form $\begin{pmatrix} u & v \\ v & -u \end{pmatrix}$; Minami estimates were obtained in [11] in the situation in which both u and v are β-regular random variables. Still, this leaves out the case where v is deterministic, which may be called a fully diluted model (meaning that there is only one degree of freedom in the randomness for each block).

In this work we consider a fully diluted random block Hamiltonian with 2×2 block $\begin{pmatrix} u & 1 \\ 1 & -u \end{pmatrix}$. This is a model originally proposed by Spencer; localization and Hölder continuity of the integrated density of states are proven in [12]. We prove a probabilistic statement on minimum eigenvalue spacing. Specifically, we give a bound showing that the probability that the minimum eigenvalue separation is less than δ tends to zero with

2

δ. (See Theorem 3.9 below). Our interest in this problem stems from recent work [21] in which a minimum eigenvalue separation condition (called limited level attraction) was used as an assumption in a proof of many-body localization for a one-dimensional spin chain. Although the separation condition remains unproven in the many-body context, one may perhaps gain some insight into the problem by replacing the tensor product of 2×2 spin spaces with a direct sum, and this leads to the random block Hamiltonian considered in this paper. The dilution of randomness is not as extreme as in the many-body case, but it remains an important issue nevertheless.

The method should be of interest in its own right. We use a sequence of Schur complements to reduce the Hilbert space to subspaces of smaller and smaller dimension. At the same time, we shrink the width of the energy interval under consideration, and eventually only one eigenvalue remains. The subspaces are structured in physical space as a set of extended blocks whose (renormalized) spectrum is in a small neighborhood of some energy E. The Schur complement formula produces effective Hamiltonians for the blocks exhibiting exponential decay in the distance between blocks. The blocks themselves become more and more dilute as the procedure proceeds, so the distance between blocks grows, and hence inter-block matrix elements tend rapidly to zero. The key probabilistic input involves estimating the probability that a block has spectrum close to E, and bounding the probability that a block has a very small level spacing. The determinant and the discriminant of the Schur complement matrix for a block is a polynomial in the random variables (the degree depends on the size of the block). This ensures a degree of transversality of eigenvalues of subsystems (and their differences) – although not uniformly in the size of the subsystem. We obtain a rather weak bound on the probability of spectrum near E (replacing the usual Wegner estimate [34]) and the probability of near-degeneracy (replacing the usual Minami estimate). Although the bound degenerates as the size of the block gets large, large blocks are improbable to begin with, and the trade-off determines the form of our results on density of states and level separation. There is some similarity with the method of [5, 26], in which localization is proven for a Hamiltonian depending analytically on the random variables. In those works, the Wegner estimate is replaced with a transversality condition arising from polynomial approximation of determinants (using a version of Cartan Lemma proven in [5]).

As in [20], we give a direct construction of eigenfunctions and eigenvalues in terms of convergent graphical expansions that depend on the random variables only within each graph. However, we gain some precision by working near a fixed energy (as advocated by [8]). We have also been influenced by ideas in [3, 2] (the Feshbach map) and of course by [1, 15].

1.2 Model

We consider the following non-monotone Anderson model. A particle hops in Λ, a rectangle in $\mathbb{Z}^d$, with each position $x \in \Lambda$ having two states, $+$ and $-$. The single-

position Hamiltonian is

$$h_x \equiv \begin{pmatrix} u_x & 1 \\ 1 & -u_x \end{pmatrix}, \tag{1.1}$$

where $\mathbf{u} = \{u_x\}_{x \in \Lambda}$ is a collection of bounded iid random variables with a bounded density. For simplicity, let us assume each u_x has a uniform distribution on [-1,1]. The Hilbert space is $\mathbb{C}^{2n}$, where $n = |\Lambda|$ is the number of elements of Λ. Let $H_0 = \oplus_{x \in \Lambda} h_x$. The hopping matrix is

$$V = \gamma J \otimes I_2, \tag{1.2}$$

with I_2 denoting the 2×2 identity matrix for the internal space, γ is a small parameter, and

$$J_{xy} = \begin{cases} 1, & \text{if } |x - y| = 1; \\ 0, & \text{otherwise.} \end{cases} \tag{1.3}$$

Here $|x| \equiv \sum_{i=1}^d x_i$. Finally, we define the Hamiltonian

$$H = H_0 + V. \tag{1.4}$$

One could of course consider a much wider class of models. However, this choice will serve to illustrate the method in a simple context exhibiting both non-monotonicity and dilution. That is, the "bare" energies $u_x, -u_x$ in (1.1) move in opposite directions as u_x varies, and there is only one random variable u_x for each 2×2 block of H_0.

1.3 Main Results

Our goal is first to prove localization by controlling the density of states and by showing stretched exponential decay of the eigenfunction correlator $\mathbb{E} \sum_\alpha |\varphi_\alpha(x)\varphi_\alpha(y)|$. Second, we prove probabilistic bounds on level spacing by constructing individual eigenstates and controlling their separation in energy from the others.

We prove several facts about the random block Hamiltonian (1.4).

Theorem 1.1. *(Theorem 3.1). There exists a constant $b > 0$ such that for γ sufficiently small the following is true. For any $0 < \delta \leq \gamma^{1/4}$, any rectangle Λ, and any interval $I = [E - \delta/2, E + \delta/2]$, let $N(I)$ denote the number of eigenvalues of H in I. Then*

$$\mathbb{E} \, N(I) \leq 4|\Lambda| \exp(-b|\log \gamma|^{2/3} |\log \delta|^{1/3}). \tag{1.5}$$

Although this bound goes to zero faster than any power of $1/|\log \delta|$, it does not imply Hölder continuity of the integrated density of states. Thus it is weaker than the result of [12].

Theorem 1.2. *(Theorem 3.7). Let $\chi = \frac{2}{9}$. There exists a constant $\bar{q} > 0$ such that for γ sufficiently small, the eigenfunction correlator satisfies*

$$\mathbb{E} \sum_\alpha |\varphi_\alpha(y)\varphi_\alpha(z)| \leq 4\gamma^{\bar{q}|y-z|^\chi}, \tag{1.6}$$

for any rectangle Λ.

4

A similar bound is proven in [12], but with exponential decay. If we restrict the sum in (1.6) to eigenvalues in I, we prove a similar bound but with a prefactor tending to zero like the right-hand side of (1.5) – Corollary 3.8. We prove these results partly as a warm-up to the level-spacing problem – but it is also helpful to understand how strong a result is possible with our method and to make comparisons with the fractional-moment results.

Finally, we give our main result on minimum level spacing. Let $\{E_\alpha\}_{\alpha=1,\ldots,2|\Lambda|}$ denote the eigenvalues of H.

Theorem 1.3. *(Theorem 3.9). There exists a constant $b > 0$ such that for γ sufficiently small,*

$$P\left(\min_{\alpha\neq\beta}|E_\alpha - E_\beta| < \delta\right) \leq |\Lambda|\exp\left(-b|\log\gamma|^{3/4}|\log\delta|^{1/4}\right), \qquad (1.7)$$

for any rectangle $|\Lambda|$ and any $0 < \delta \leq \gamma^{1/4}/4$.

As this is a result about microscopic level statistics, we are interested in the behavior as $\delta \to 0$. This bound tends to zero more slowly than (1.5), but still faster than any power of $1/|\log\delta|$.

1.4 Fundamental Lemma

We state here a lemma on Schur complements that will be used often in our analysis.

Lemma 1.4. *(Fundamental Lemma) Let K be a $(p+q) \times (p+q)$ symmetric matrix in block form, $K = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$, with A a $p \times p$ matrix, D a $q \times q$ matrix, and $C = B^T$. Assume that $\|(D-E)^{-1}\| \leq \varepsilon^{-1}, \|B\| \leq \gamma, \|C\| \leq \gamma$. Define the Schur complement with respect to λ:*

$$F_\lambda \equiv A - B(D-\lambda)^{-1}C. \qquad (1.8)$$

Let ε and γ/ε be small, and $|\lambda - E| \leq \varepsilon/2$. Then

1. *If φ is an eigenvector for F_λ with eigenvalue λ, then $(\varphi, -(D-\lambda)^{-1}C\varphi)$ is an eigenvector for K with eigenvalue λ, and all eigenvectors of K with eigenvalue λ are of this form.*

2. *The spectrum of K in $[E - \varepsilon/2, E + \varepsilon/2]$ is in close agreement with that of F_E in the following sense. If $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_m$ are the eigenvalues of K in $[E - \varepsilon/2, E + \varepsilon/2]$, then there are corresponding eigenvalues $\tilde\lambda_1 \leq \tilde\lambda_2 \leq \ldots \leq \tilde\lambda_m$ of F_E, and $|\lambda_i - \tilde\lambda_i| \leq 2(\gamma/\varepsilon)^2|\lambda_i - E|$.*

Note that the degree of closeness between the two sets of eigenvalues improves the closer one gets to $\lambda = E$.

Our goal, then, is to iterate the fundamental lemma, using appropriately chosen subspaces, so that the spectral window width ε tends to zero, with γ/ε tending to zero as well. Eventually, the width will be so small that at most one eigenvalue will be

inside. At that point, the decay of the eigenfunction will become manifest, as will the separation of the eigenvalue from the rest of the spectrum.

The operator $-(D - \lambda)^{-1}C$ plays the role of expanding a solution of a low dimensional space to a higher dimensional space. Such factors will be composed with each other to obtain the proper eigenfunction on the full Hilbert space; hence we call them *eigenfunction-generating kernels*.

Proof. The first statement is a standard fact about the Schur complement. If φ is an eigenvector for F_λ with eigenvalue λ, then

$$(K - \lambda) \begin{pmatrix} \varphi \\ -(D - \lambda)^{-1}C\varphi \end{pmatrix} = \begin{pmatrix} (A - \lambda)\varphi - B(D - \lambda)^{-1}C\varphi \\ C\varphi - (D - \lambda)(D - \lambda)^{-1}C\varphi \end{pmatrix}$$
$$= \begin{pmatrix} (F_\lambda - \lambda)\varphi \\ 0 \end{pmatrix} = 0. \tag{1.9}$$

Conversely, if $(K - \lambda)(\varphi, \tilde{\varphi}) = 0$, then $C\varphi + (D - \lambda)\tilde{\varphi} = 0$, which has unique solution $\tilde{\varphi} = -(D - \lambda)^{-1}C\varphi$. Substituting in, we obtain $(F_\lambda - \lambda)\varphi = 0$. Note that this implies agreement between the multiplicities of the λ-eigenspaces for H and F_λ.

For the second statement, write

$$F_E - F_\lambda = -B(D - E)^{-1}C + B(D - \lambda)^{-1}C = B(D - E)^{-1}(\lambda - E)(D - \lambda)^{-1}C. \tag{1.10}$$

Noting that $|\lambda - E| \leq \varepsilon/2$ and $\mathrm{dist}(\mathrm{spec}\, D, E) \geq \varepsilon$, we have that $\|(D - \lambda)^{-1}\| \leq 2/\varepsilon$, and thus

$$\|F_E - F_\lambda\| \leq \frac{2\gamma^2}{\varepsilon^2}|\lambda - E|. \tag{1.11}$$

By Weyl's inequality, the eigenvalues of F_λ and F_E differ by no more than $2(\gamma/\varepsilon)^2|\lambda_i - E|$ when shifting from F_λ to $F_E = F_\lambda + (F_E - F_\lambda)$. $\square$

2 Fixed Energy Procedure

2.1 First Step

Let us investigate the spectrum of H in the vicinity of E. The energy E can be any real number (although since H is bounded, there will be no spectrum near E if $|E|$ is large). For the time being, E will be treated as a fixed parameter. We emphasize that we are deriving probabilistic bounds on resonances by way of the geometric/algebraic behavior of the eigenvalues with respect to the random parameters. Let us observe this in the simplest possible case.

It will be helpful to introduce some terminology. The rectangle Λ is a collection of *positions* in $\mathbb{Z}^d$ labeling the coordinates in the first factor in $\mathcal{H} = \mathbb{C}^{|\Lambda|} \otimes \mathbb{C}^2$. Coordinates for the full $\mathcal{H}$ will be called *sites*; thus there are $|\Lambda|$ positions and $2|\Lambda|$ sites. Let Λ^* denote the set of sites. More generally, let X^* denote the set of sites at positions $x \in X \subset \Lambda$. It can be visualized as two copies of Λ in two layers: two sites at each position.

The eigenvalues of h_x are $\pm t_x$, where

$$t_x = \sqrt{u_x^2 + 1}. \tag{2.1}$$

We say that x is ε-resonant with E if either eigenvalue is within ε of E. Sites associated with non-resonant positions are also said to be non-resonant. The eigenvalues are split by at least 1, hence for ε small only one can be close to E. A direct calculation shows that

$$P(t_x \in [|E| - \varepsilon, |E| + \varepsilon]) \le P(t_x \in [1, 1 + 2\varepsilon]) = 2\sqrt{\varepsilon + \varepsilon^2} \le 3\sqrt{\varepsilon}, \text{ for } \varepsilon \le 1, \tag{2.2}$$

assuming a uniform distribution on [-1,1] for u_x. (The maximum is achieved when $|E| = 1 + \varepsilon$.) Thus $3\sqrt{\varepsilon}$ is a bound for the probability that x is an ε-resonant position with respect to E. We take $\varepsilon = \gamma^\phi$ with $\phi = \frac{1}{4}$.

Let $R^{(1)}$ denote the set of ε-resonant positions $x \in \Lambda$. It can be broken into components, where connectedness is based on nearest-neighbor connections. Let us estimate $\mathcal{P}_{x,y}^{(1)}$, the probability that x and y are in the same component. For positions x and y to be connected, there must be a self-avoiding walk from x to y, consisting of nearest-neighbor steps between resonant sites. By (2.2), the probability that all the sites in a given walk are resonant is bounded by $(3\sqrt{\varepsilon})^m$, where m is the number of positions in the walk, including endpoints. If we add a combinatoric factor $2d$ per step, we may convert the sum over walks with m positions into a supremum[1]. A combinatoric factor 2^m allows us to fix m. We obtain

$$\mathcal{P}_{x,y}^{(1)} \le (12d\sqrt{\varepsilon})^m \le (12d\sqrt{\varepsilon})^{|x-y|+1}, \tag{2.3}$$

as $m - 1$ must be at least $|x - y|$, the minimum number of steps to get from x to y.

Block form of the Hamiltonian. Recall that $R^{(1)}$ is the set of resonant positions. Putting $R^{(1)c} = \Lambda \setminus R^{(1)}$, we have associated sets of sites $R^{(1)*}$, $R^{(1)c*}$. These index sets determine the block form of the Hamiltonian:

$$H = \begin{pmatrix} A^{(1)} & B^{(1)} \\ C^{(1)} & D^{(1)} \end{pmatrix}, \tag{2.4}$$

with $A^{(1)}$ representing the restriction of H to $R^{(1)*}$, and $D^{(1)}$ representing the restriction to $R^{(1)c*}$ (or, to be precise, these are the restrictions of H to the subspaces corresponding to those index sets). Note that $A^{(1)}, D^{(1)}$ contain off-diagonals only between sites at the same position or at adjacent positions; positions which belong to different components in $R^{(1)}$ do not interact.

We may write

$$D^{(1)} = W^{(1)} + V^{(1)}, \tag{2.5}$$

with $W^{(1)}$ block diagonal (with blocks h_x at each position x) and with $V^{(1)}$ equal to the projection of V to $R^{(1)c}$. Let $V_{xy}^{(1)}$ denote the block matrix element of $V^{(1)}$ between

---

[1] Combinatoric factors are a convenient bookkeeping device when estimating sums. If $\sum_\alpha c_\alpha^{-1} \le 1$ for some positive constants c_α, then $\sum_\alpha X_\alpha \le \sup_\alpha |X_\alpha| c_\alpha$. We call c_α a combinatoric factor.

positions x and y. For $x, y \in R^{(1)\mathrm{c}}$, it is equal to $\gamma J_{xy} I_2$, see (1.2). Thus it is nonzero only for adjacent positions x and y in $R^{(1)\mathrm{c}}$, in which case it equals γI_2.

Convention. We use ordinary absolute value signs $|\cdot|$ to denote the norm of a 2×2 matrix such as the block matrix element $V_{xy}^{(1)}$. This is to distinguish it from the usual notation $\|\cdot\|$ for matrices indexed by Λ^* or some subset of Λ^*. Using this notation, we have that $|V_{xy}^{(1)}| = \gamma$ for x, y adjacent positions in $R^{(1)\mathrm{c}}$.

Random walk expansion. We would like to reduce (2.4) by Schur complementation:

$$F_\lambda^{(1)} \equiv A^{(1)} - B^{(1)}(D^{(1)} - \lambda)^{-1} C^{(1)}. \tag{2.6}$$

Let us develop a random walk expansion for $(D^{(1)} - \lambda)^{-1}$.

By construction, $\operatorname{dist}(\operatorname{spec} W^{(1)}, E) \geq \varepsilon$. Let us assume that $|\lambda - E| \leq \varepsilon/2$, so that $\|(W^{(1)} - \lambda)^{-1}\| \leq 2/\varepsilon$. The off-diagonal matrix $V^{(1)}$ has norm no greater than $2d\gamma$ in dimension d (it is a symmetric matrix with row sum equal to $2d\gamma$). Therefore, for $\varepsilon = \gamma^\phi = \gamma^{1/4}$ sufficiently small, $(D^{(1)} - \lambda)^{-1}$ exists and can be expanded in a Neumann series:

$$(D^{(1)} - \lambda)^{-1} = (W^{(1)} - \lambda)^{-1} + (W^{(1)} - \lambda)^{-1} V^{(1)} (W^{(1)} - \lambda)^{-1} + \cdots. \tag{2.7}$$

This leads to a random walk expansion with steps determined by block matrix elements of $V^{(1)}$:

$$\left[(D^{(1)} - \lambda)^{-1}\right]_{xy} = \sum_{g_1 : x \to y} \prod_{i=1}^{m} [(W^{(1)} - \lambda)^{-1}]_{x_i x_i} \prod_{j=1}^{m-1} V_{x_j x_{j+1}}^{(1)}. \tag{2.8}$$

Here $g_1 = \{x = x_1, x_2, \ldots, x_m = y\}$ is a random walk with each x_i, x_{i+1} adjacent positions; return visits are allowed. For such a walk we write its length as $|g_1| = m$. Note that each factor on the right-hand side of (2.8) is a 2×2 block matrix element. As each eigenvalue of $W^{(1)} - \lambda$ is at least $\varepsilon/2$ in magnitude, we have that $|[(W^{(1)} - \lambda)^{-1}]_{x_i x_i}| \leq 2/\varepsilon$ for every $x_i \in g_1$. Thus a term in the sum (2.8) with $|g_1| = m$ is bounded by $\gamma^m (\varepsilon/2)^{-(m+1)}$. We may assign combinatoric factors 2^m to fix the length of the walk and $2d$ for each sum over x_j; altogether the combinatoric factor is $2^m (2d)^{m-1}$. Thus we obtain a bound

$$\left| \left[(D^{(1)} - \lambda)^{-1}\right]_{xy} \right| \leq \left(\frac{c_d \gamma}{\varepsilon}\right)^{|x-y|} \cdot \frac{4}{\varepsilon}, \tag{2.9}$$

as $m - 1$ must be at least $|x - y|$, the number of steps needed to walk from x to y. Here, and in what follows, c_d denotes a constant that depends only on the dimension d.

Observe from (2.7) and (2.8) that the operator $(D^{(1)} - \lambda)^{-1}$ connects all sites in $R^{(1)\mathrm{c}*}$. Therefore, we see that resonant sites $R^{(1)*}$ are connected through the chain of operators $B^{(1)}(D^{(1)} - \lambda)^{-1} C^{(1)}$ in the second term of (2.6). Note that $B^{(1)}$ and $C^{(1)}$ contain off-diagonal interactions from V. Thus the walks in (2.8) are extended with two additional steps $V_{x_j x_{j+1}}^{(1)}$. As in the arguments for (2.9), we obtain

$$\left| \left[B^{(1)}(D^{(1)} - \lambda)^{-1} C^{(1)}\right]_{xy} \right| \leq \left(\frac{c_d \gamma}{\varepsilon}\right)^{|x-y| \vee 2} \cdot \varepsilon. \tag{2.10}$$

8

In a similar manner, we can estimate the *eigenfunction-generating kernel* for the first step:

$$\left| \left[(D^{(1)} - \lambda)^{-1} C^{(1)} \right]_{xy} \right| \leq \left(\frac{c_d \gamma}{\varepsilon} \right)^{|x-y| \vee 1}. \tag{2.11}$$

The graphical expansion (2.8) induces similar expansions for these matrix elements.

2.2 Resonant Blocks

In the kth step we use a length scale $L_k \equiv 2^{k-1}$ to determine connectivity. That is, in the first step, $L_1 = 1$ and two positions are in the same component if they are nearest neighbors. In steps $k \geq 2$, we inherit a set of resonant positions $R^{(k-1)}$, and two positions are in the same component if they are within a distance L_k. Thus, components at scale L_k are subsets of $R^{(k-1)}$ that can be spanned by connections of length up to L_k. A component of cardinality n has $2n$ sites; each of these is effectively connected to all the others of the component. Distance will always be measured in terms of the original metric $|x| = \sum_{i=1}^{d} |x_i|$ on $\mathbb{Z}^d$.

Definition 2.1. *Let B be a component of $R^{(k-1)}$ on scale k with $k \geq 2$. We say that B is **isolated on scale k** if*

  *1. its **volume** $(\equiv n(B) \equiv$ number of positions) is no greater than $L_k^{2\psi/5}$,*

  2. its distance from every other component on scale k is at least $4L_k$,

  3. its diameter is $\leq L_k$.

Here $\psi = \frac{2}{3}$ is an exponent governing fractional exponential decay of graphical bounds.

From the previous step, we have the Schur complement matrix

$$F_\lambda^{(k-1)} = A^{(k-1)} - B^{(k-1)} (D^{(k-1)} - \lambda)^{-1} C^{(k-1)}, \tag{2.12}$$

with bounds similar to (2.10) that are proven below in Theorem 2.7 (and that can be assumed by induction). It operates on the space indexed by elements of $R^{(k-1)*}$. Each of the matrices in (2.12) is given by a random walk expansion as in (2.8) in terms of the corresponding matrices from step $k-2$. See Section 2.4 for details on the random walk expansion; for now, we need only some of its general properties.

In each step we take the Schur complement of the current renormalized Hamiltonian. In the first step, $F^{(1)}$ is a renormalization of the original Hamiltonian H; the interactions of the first step generate walks, or *graphs* constructed of connections in the original Hamiltonian. Further renormalizations $F^{(2)}, F^{(3)}, \ldots$ involve walks of the preceding Hamiltonian's walks; we will refer to these nested walk structures as *multigraphs*. Thus the Schur complement $F_\lambda^{(k-1)}$ is defined inductively in terms of $(k-1)$-level *multigraphs*, each being a walk with steps given by $(k-2)$-level multigraphs, and so on down to the first level. By construction, these graphs sample the potential u_x only for positions x that the first-level walk passes through; thus there is dependence on u_x only for

$x \notin R^{(k-1)}$. Consequently, the dependence on u_x, $x \in R^{(k-1)}$ is only through the leading term in $F_\lambda^{(k-1)}$, specifically the matrix $A^{(k-1)}$, projected to $R^{(k-1)*}$. (Recall that $A^{(1)}$ is the original Hamiltonian H projected to $R^{(1)*}$.) In particular, the values $u_x \in R^{(k-1)}$ *do not* appear in any of the block inverse operators used – see (2.8) and its kth step version (2.55).

One may visualize the last term in (2.12) as a set of graphs that exit $R^{(k-1)}$ (the matrix $B^{(k-1)}$), and then wander in $R^{(k-2)} \setminus R^{(k-1)}$ (the matrix $(D^{(k-1)} - \lambda)^{-1}$) before returning to $R^{(k-1)}$ (the matrix $C^{(k-1)}$). Thus we see that $F_\lambda^{(k-1)}$ agrees with H on $R^{(k-1)*}$ up to small terms from the random walk expansions, all independent of u_x, $x \in R^{(k-1)}$. Hence $F_\lambda^{(k-1)}$ is equal to a direct sum of H projected to the $R^{(k-1)}$ components plus small $O(\gamma^2/\varepsilon)$ corrections (independent of u_x, $x \in R^{(k-1)}$). That is, we may write

$$F_\lambda^{(k-1)} = \oplus_B h_B + \Xi_{k-1}, \tag{2.13}$$

where h_B is the projection of H to the component B in $R^{(k-1)}$ and $\Xi_{k-1} = O(\gamma^2/\varepsilon)$ is independent of u_x, $x \in R^{(k-1)}$. For $k = 2$ (2.13) holds by (2.10) and indeed we will generalize it for $k > 2$ in the induction procedure – see Corollary 2.8. Furthermore, there is no off-diagonal dependence on u_x, $x \in R^{(k-1)}$.

We wish to show that isolated components are unlikely to contain spectrum near E. We could work with $F_\lambda^{(k-1)}$, projected to an isolated component B of $R^{(k-1)}$, but in order to preserve independence we truncate the random walk expansions in $F_\lambda^{(k-1)}$, including only multigraphs that remain within a distance $< L_{k-2}$ of B. Multigraphs neglected therefore travel a distance $\geq 2L_{k-2} = L_{k-1}$ after returning to B or reaching another block. Let us designate such multigraphs *nonlocal* at scale L_{k-1}; they visit points at a distance $\geq L_{k-2}$ from B. Let us write $\tilde{F}_E^{(k-1)}$ for the renormalized Hamiltonian with all nonlocal multigraphs dropped. The truncated Hamiltonian is block diagonal due to the separation condition: $\tilde{F}_E^{(k-1)} = \oplus_B \tilde{F}_E^{(k-1)}(B)$. As above, we may write $\tilde{F}_\lambda^{(k-1)}(B) = h_B + \tilde{\Xi}_{k-1}(B)$ for the truncated $F_\lambda^{(k-1)}$, projected to block B, and $\tilde{\Xi}_{k-1}(B) = O(\gamma^2/\varepsilon)$.

We prove decay at rate γ^{L^ψ} for graphs of length L, with $\psi = \frac{2}{3}$ – see Theorem 2.7 below. We prove below a bound on the probability of spectrum within $\varepsilon^{L_{k-1}^\psi}$ of E. As this window is much larger than the size of terms neglected (recall that $\varepsilon = \gamma^\phi$ with $\phi = \frac{1}{4}$), movement of the spectrum due to terms of order $\gamma^{L_{k-1}^\psi}$ is insignificant.

Definition 2.2. *Let B be a component of $R^{(k-1)}$ on scale k with $k \geq 2$. We say that B is **resonant on scale k** if it is isolated on scale k and if $\mathrm{dist}(\mathrm{spec}\, \tilde{F}_E^{(k-1)}(B), E) \leq \varepsilon_k \equiv \varepsilon^{L_k^\psi}$. Here we refer to the truncation $F_E^{(k-1)} \to \oplus_B \tilde{F}_E^{(k-1)}(B)$, which results from the deletion of all multigraphs that extend a distance L_{k-2} or farther from any block B.*

Following a similar plan as in the first step, we define the new resonant set $R^{(k)}$ by removing from $R^{(k-1)}$ all the non-resonant isolated components. Thus

$$R^{(k)} = R^{(k-1)} \setminus \bigcup_{\alpha:\, B_\alpha \text{ is isolated but not resonant on scale } k} B_\alpha. \tag{2.14}$$

10

Proposition 2.3. *Let B be an isolated component of $R^{(k-1)}$ on scale k with $k \geq 2$. The probability that B is resonant on scale k is less than $\varepsilon^{L_k^{3\psi/5}/3}$.*

Proof. Let us enumerate the positions in B as $\{x_1, \ldots, x_n\}$. The block B is connected on scale L_{k-1} (*i.e.* using connections between x_i and x_j of length up to L_{k-1}. Let us consider the $2n \times 2n$ matrix $\tilde{F}_\lambda^{(k-1)}(B)$ as a function of $\mathbf{u} = \{u_{x_1}, \ldots, u_{x_n}\}$, with all other u's held fixed, which means that $\tilde{\Xi}_{k-1}(B)$ is fixed as well. Thus, as explained above, $\tilde{F}_\lambda^{(k-1)}(B)$ agrees with H up to $\mathbf{u}$-independent terms of size $O(\gamma^2/\varepsilon)$. We consider here $\lambda = E$, since by the fundamental lemma, control of the spectrum of $\tilde{F}_E^{(k-1)}(B)$ allows us to control $\tilde{F}_\lambda^{(k-1)}(B)$ for λ near E. Recall that H has u_i or $-u_i$ on the diagonal, and it has off-diagonal matrix elements connecting adjacent positions and connecting the two sites at each position. The off-diagonal matrix entries are either γ, 1, or 0. It should be clear, then, that $\Delta(\mathbf{u}) \equiv \det(\tilde{F}_E^{(k-1)}(B) - E)$ is a polynomial in $\mathbf{u}$ of degree $2n$. We may therefore apply standard results bounding the measure of sets where the polynomial is small.

Noting that each $u_x \in [-1, 1]$, and since $t_x = \sqrt{u_x^2 + 1}$, the spectrum of H is clearly limited to the range $[-\sqrt{2}, \sqrt{2}] + O(\gamma)$. Therefore, we can assume $E \in [-\frac{3}{2}, \frac{3}{2}]$; this will cover all of the spectrum of H.

With $E \in [-\frac{3}{2}, \frac{3}{2}]$, we need to establish in a quantitative sense that $\Delta(\mathbf{u}) = \det(\tilde{F}_E^{(k-1)}(B) - E)$ is not the zero polynomial. Thus we look for a value of $\mathbf{u}$ for which $\Delta(\mathbf{u})$ is bounded away from 0. Take $\mathbf{u}_0 = (2, \ldots, 2)$. Then each $t_{x_i} = \sqrt{5}$, and so

$$(t_x - E)(-t_x - E) = E^2 - 5 \leq \tfrac{9}{4} - 5 = -\tfrac{11}{4}. \tag{2.15}$$

We must allow for $O(\gamma)$ movement of the eigenvalues due to the terms other than $\oplus_i h_{x_i}$ in $\tilde{F}_E^{(k-1)}(B)$ (here we use also the decay away from the diagonal – see Theorem 2.7 below – and the fact that the norm of the perturbing matrix is bounded by the maximum absolute row sum). Still, we maintain a lower bound $|\Delta(\mathbf{u}_0)| \geq 2^n$, after taking the product over the n 2×2 blocks and allowing for the small changes in the eigenvalues.

We use the Brudnyi-Ganzburg inequality, in the form stated in [6]:

$$\sup_U |p| \leq \left(\frac{4n|U|}{|\omega|}\right)^\kappa \sup_\omega |p|. \tag{2.16}$$

Here U is a bounded convex subset of $\mathbb{R}^n$, p is a polynomial of degree at most κ and ω is a measurable subset of U. In our case, we let $U = [-2, 2]^n, \kappa = 2n$, and

$$\omega = \{\mathbf{u} : |\Delta(\mathbf{u})| \leq \varepsilon_k 3^{2n}\} \text{ with } \varepsilon_k \equiv \varepsilon^{L_k^\psi}. \tag{2.17}$$

We have shown that $\sup_U |\Delta(\mathbf{u})| \geq 2^n$ and so

$$|\omega| \leq \left(2^{-n} \varepsilon_k 3^{2n}\right)^{\frac{1}{2n}} \cdot 4n|U| \leq \varepsilon_k^{1/(2n)} 6\sqrt{2} \cdot n 4^n. \tag{2.18}$$

11

Since u_x is distributed uniformly on [-1,1], this leads to bounds

$$P(|\Delta(\mathbf{u})| \leq \varepsilon_k 3^{2n}) \leq \varepsilon^{L_k^\psi/(2n)} \cdot 9n2^n$$
$$\leq \varepsilon^{L_k^{3\psi/5}/2} \cdot 9L_k^{2\psi/5} 2^{L_k^{2\psi/5}} \leq \varepsilon^{L_k^{3\psi/5}/3}, \qquad (2.19)$$

where in the second line we have used the bound $n \leq L_k^{2\psi/5}$ from Definition 2.1 (as well as the smallness of $\varepsilon = \gamma^\phi$). Note that both E and the eigenvalues of $\tilde{F}_E^{(k-1)}(B)$ are in $[-\frac{3}{2}, \frac{3}{2}]$, and therefore the eigenvalues of $\tilde{F}_E^{(k-1)}(B) - E$ are no greater than 3 in magnitude. Hence (2.19) implies the statement of the proposition. $\qquad \square$

2.3 Probability Estimates

We need to work on a multiscale version of the percolation estimate (2.3). As we investigate the percolation of resonant blocks, we face a situation in which blocks of $R^{(k)}$ may inherit probability bounds from step $k-1$ or they may receive a probability bound from Proposition 2.3. Thus we make an inductive definition that keeps track of the probability bounds available for blocks at any scale:

$$\tilde{P}^{(1)}(B) \equiv (3\sqrt{\varepsilon})^{n(B)}$$
$$\tilde{P}^{(k)}(B) \equiv \begin{cases} \varepsilon^{L_{k-2}^{3\psi/5}/3}, & \text{if } B \text{ is isolated on scale } k-2; \\ \prod_{i=1}^{m'} \tilde{P}^{(k-1)}(B_i), & \text{otherwise.} \end{cases} \qquad (2.20)$$

We emphasize that the $\tilde{P}$ are not themselves probabilities but stand in for upper bounds on the probabilities of resonant clusters. Here $n(B)$ is the number of positions in B, and $B_1, \ldots, B_{m'}$ are the subcomponents of B on scale $k-1$. Note that each factor $3\sqrt{\varepsilon}$ or $\varepsilon^{L_{k-2}^{3\psi/5}/3}$ (corresponding to a bound on the probability that a site or block is resonant) carries forward to the next scale, until a scale k is reached with B isolated on scale $k-2$. In that case, all the factors from subblocks of B from previous scales are replaced with the bound from Proposition 2.3. This corresponds to replacement of bounds by a better one from Proposition 2.3, when it is available (*i.e.* if a component is isolated on scale $k-2$).

Next let us define a weighted sum of the probability bounds $\tilde{P}^{(k)}(B)$:

$$Q_{x,y}^{(k)} \equiv \sum_{B \text{ containing } x \text{ and } y} \tilde{P}^{(k)}(B)\varepsilon^{-q_k n(B)}. \qquad (2.21)$$

Here $\{q_k\}$ is a decreasing sequence of numbers with a strictly positive floor $q > 0$; it will be specified in the next theorem. We prove the following "energy-entropy" bound, similar in spirit to [15]. It is a purely geometric-combinatoric result about the components B that are generated by our definitions.

Theorem 2.4. *Given $\psi = \frac{2}{3}$, let $\chi = \psi/3$. Then for ε sufficiently small, and $q_1 \equiv \frac{1}{5}$, $q_2 \equiv \frac{1}{6}$, $q_3 \equiv \frac{1}{7}$, and $q_k \equiv q_{k-1}(1 - L_{k-2}^{-\psi/20})$ for $k \geq 4$, we have*

$$Q_{x,y}^{(k)} \leq \varepsilon^{q_k(|x-y| \vee L_{k-2})^\chi}. \qquad (2.22)$$

Proof. To treat the case $k = 1$, we need to extend the proof of (2.3) to incorporate decay with $n = n(B)$, the volume of the component containing x, y. The walk from x to y can be extended to a branching random walk (*i.e.* a spanning tree). The number of such trees is bounded by c_d^n. The weighting $\varepsilon^{-q_1 n} = \varepsilon^{-n/5}$ in (2.21) is counterbalanced by our bound on the probability that n positions are resonant, $(3\sqrt{\varepsilon})^n$. Noting that $n \geq |x - y| + 1$, we may extract the decay in (2.22), with a decrease of an additional $1/5$ in the power of ε. This leaves a bound of $(2c_d\varepsilon^{1/10})^n$. The sum on n is less than 1, and hence (2.22) holds for $k = 1$.

Let us now assume (2.22) for $k - 1$. Let us write

$$Q_{x,y}^{(k)} = Q_{x,y}^{(k),1} + Q_{x,y}^{(k),2}, \tag{2.23}$$

with the first term giving the sum over blocks B that are *not* isolated on scale $k - 2$, and the second term giving the sum over the ones that *are* isolated. We consider first $Q_{x,y}^{(k),1}$, and note that $\tilde{P}^{(k)}(B)$ is given by the second alternative in (2.20). Recall that B is a component of $R^{(k)}$ and that it is formed by joining together components of $R^{(k-1)}$ with links of length $\leq L_k$. Hence there must be a sequence of positions $x = x_1, y_1, x_2, y_2, \ldots, x_m, y_m = y$ such that each x_i, y_i lie in the same component B_i of $R^{(k-1)}$, and $|x_i - y_{i-1}| \leq L_k$. The remaining subcomponents of B may be denoted $B_{m+1}, \ldots, B_{m'}$. Connectivity on scale k implies that there must be a tree graph T connecting all the components of B. Each link of the tree graph is a pair (x_i, y_{i-1}) as above with $2 \leq i \leq m$ or a pair $(\tilde{x}_j, \tilde{y}_j)$ with $m + 1 \leq j \leq m'$, $\tilde{x}_j \in B_j$, $\tilde{y}_j \in B_{\tau(j)}$, and $\tau(j) < j$. Thus we choose a tree graph that extends the unbranched connection from B_1 to B_m. (See Figure 1.) Again, because of connectivity, we can assume that $|\tilde{x}_j - \tilde{y}_j| \leq L_k$.

We may write

$$Q_{x,y}^{(k),1} \leq \sum_{m,m'} \sum_{T} \sum_{\substack{B_1,\ldots,B_{m'} \\ \text{consistent with } T}} \prod_{i=1}^{m'} \left[\tilde{P}^{(k-1)}(B_i)\varepsilon^{-q_k n(B_i)} \right], \tag{2.24}$$

where we have written $n = \sum_i n(B_i)$ as the sum of the volumes of the B_i. With some overcounting, we may proceed through the tree graph, summing over vertices $x_i, y_i, \tilde{x}_j, \tilde{y}_j$ and components, requiring only that each B_i contain a specified set of 1 or 2 vertices. In this way, we will be able to apply the inductive hypothesis from step $k - 1$.

The tree graph T may be broken up into the chain connecting $B_1, \ldots, B_m$ plus individual tree graphs $T_1, \ldots, T_m$ with roots at the corresponding blocks. We now focus on a single B_i (which we take as fixed) and sum over T_i and the associated blocks (other than B_i) that are linked by T_i. Let us define for an arbitrary block $\bar{B}$ on scale $k - 1$:

$$K^{(p)}(\bar{B}) \equiv \sum_{\bar{T}:\, \text{depth}(\bar{T}) \leq p} \sum_{\substack{\bar{B}_1,\ldots,\bar{B}_\ell \\ \text{consistent with } \bar{T}}} \prod_{i=1}^{\ell} \left[\tilde{P}^{(k-1)}(\bar{B}_i)\varepsilon^{-q_k n(\bar{B}_i)} \right]. \tag{2.25}$$
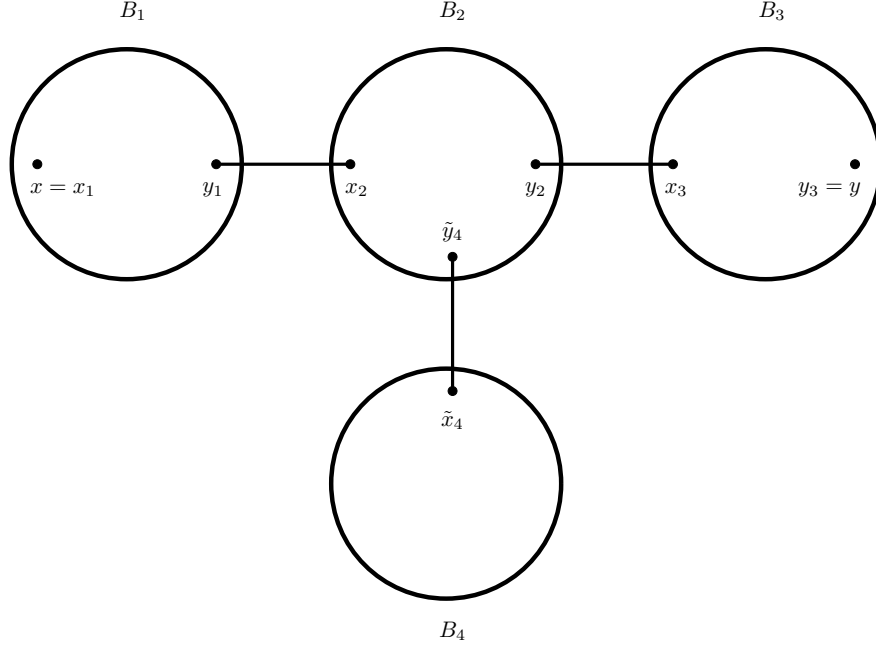
13

Figure 1: A tree graph of blocks connecting x and y. Blocks B_1, B_2, B_3 form the spine and B_4 is a leaf.

Here $\text{depth}(\bar{T})$ is the largest number of links in $\bar{T}$ that are needed to reach a vertex, starting at the root. Then we can write

$$Q_{x,y}^{(k),1} \leq \sum_{m=1}^{\infty} \sum_{y_1,x_2,y_2,\ldots,y_{m-1},x_m} \prod_{i=1}^{m} \left[\sum_{B_i \text{ containing } x_i,y_i} \tilde{P}^{(k-1)}(B_i)\varepsilon^{-q_k n(B_i)} K^{(\infty)}(B_i) \right]. \quad (2.26)$$

Here the inequality results from the overcounting associated with treating each of the T_i as independent of the others. Let us put

$$\hat{\varepsilon} = \varepsilon^{q_{k-1}L_{k-3}^{\chi}/2}, \, k \geq 2. \quad (2.27)$$

Lemma 2.5. *With* $\chi = \frac{2}{9}$ *and* ε *sufficiently small,*

$$K^{(p)}(\bar{B}) \leq e^{\hat{\varepsilon} n(\bar{B})}. \quad (2.28)$$

Proof. We use induction on p. When $p = 0$, we have the tree with no links, so (2.28) amounts to the statement $1 \leq e^{\hat{\varepsilon} n(\bar{B})}$. For larger p, we may use the recursion

$$K^{(p)}(\bar{B}) \leq \sum_{r=0}^{\infty} \frac{1}{r!} \prod_{i=0}^{r} \left[\sum_{\bar{B}_i \text{ linked to } \bar{B}} \tilde{P}^{(k-1)}(\bar{B}_i)\varepsilon^{-q_k n(\bar{B}_i)} K^{(p-1)}(\bar{B}_i) \right]. \quad (2.29)$$

Again, this is an upper bound, as we are treating each of the r branches of $\bar{T}$ as independent of the others. There is also overcounting due to permutations of the branches, and this has been compensated by the factor $1/r!$.

14

The sum over $\bar{B}_i$ linked to $\bar{B}$ may be bounded as follows. By connectivity, there must be a pair of points $\bar{x}$, $\bar{y}$ with $\bar{y} \in \bar{B}$ and $\bar{x} \in \bar{B}_i$ and $|\bar{x} - \bar{y}| \leq L_k$. The sum over $\bar{y}$ leads to a factor $n(\bar{B})$. There are no more than $(2L_k + 1)^d$ choices for $\bar{x}$, given $\bar{y}$. Furthermore, we may bound

$$\varepsilon^{-q_k n(\bar{B}_i)} K^{(p-1)}(\bar{B}_i) \leq \varepsilon^{-q_{k-1} n(\bar{B}_i)} \tag{2.30}$$

using (2.28), since

$$-(q_k - q_{k-1})|\log \varepsilon| + \hat{\varepsilon} \geq \begin{cases} |\log \varepsilon| q_{k-1} L_{k-2}^{-\psi/15}, & k \geq 4 \\ 0, & k = 2, 3 \end{cases}$$
$$\geq 0 \tag{2.31}$$

for ε small. (This follows from the expressions for q_k in Theorem 2.4, in particular the relation $q_{k-1} - q_k = q_{k-1} L_{k-2}^{-\psi/20}$ for $k \geq 4$; the increase from $\psi/20$ to $\psi/15$ is to accommodate the $\hat{\varepsilon}$ term on the left-hand side.) Inserting (2.30) into (2.29), we obtain

$$K^{(p)}(\bar{B}) \leq \exp\left((2L_k + 1)^d Q_{\bar{x},\bar{x}}^{(k-1)} n(\bar{B})\right)$$
$$\leq \exp\left((2L_k + 1)^d \varepsilon^{q_{k-1} L_{k-3}^\chi} n(\bar{B})\right) \leq e^{\hat{\varepsilon} n(\bar{B})}, \tag{2.32}$$

after applying the induction hypothesis (2.22) to $Q_{\bar{x},\bar{x}}^{(k-1)}$. This closes the induction and completes the proof of the lemma. This "polymer expansion" construction is a standard way to treat tree graph sums when the vertices have spatial extent, see for example [18]. $\qquad\square$

After using the lemma to sum over all the branches of the tree, we are left with the "spine," that is, the blocks $B_1, \ldots, B_m$ that lead from x to y. From (2.26), we obtain the following bounds:

$$Q_{x,y}^{(k),1} \leq \sum_{m=1}^{\infty} \sum_{y_1, x_2, y_2, \ldots, y_{m-1}, x_m} \prod_{i=1}^{m} \left[\sum_{B_i \text{ containing } x_i, y_i} \tilde{P}^{(k-1)}(B_i) \varepsilon^{-q_k n(B_i)} e^{\hat{\varepsilon} n(B_i)} \right] \tag{2.33}$$

$$\leq \sum_{B_1 \text{ containing } x \text{ and } y} \tilde{P}^{(k-1)}(B_1) \varepsilon^{-q_k n(B_1)} e^{\hat{\varepsilon} n(B_1)} + \sum_{m \geq 2} \sum_{y_1, x_2, y_2, \ldots, y_{m-1}, x_m} \prod_{i=1}^{m} Q_{x_i, y_i}^{(k-1)},$$

The effect of the tree graphs has been subsumed into factors $e^{\hat{\varepsilon} n(B_i)}$, by Lemma 2.5. The first term of the final expression bounds the case $m = 1$. For the terms $m \geq 2$, we have used the fact that

$$\sum_{B_i \text{ containing } x_i, y_i} \tilde{P}^{(k-1)}(B_i) \varepsilon^{-q_k n(B_i)} e^{\hat{\varepsilon} n(B_i)} \leq Q_{x_i, y_i}^{(k-1)}, \tag{2.34}$$

which follows from (2.21),(2.31).

Consider the term $m = 2$, and let $\tilde{\varepsilon} \equiv \varepsilon^{q_{k-1}}$. We need to be careful with the sums on y_1, x_2 in order to preserve the form of the decay estimate (2.22). With x, y fixed, note that

$$\sum_{y_1, x_2} \min\{|x - y_1|, |x_2 - y|\}^{-(d+1)} \leq c_d L_k^d. \tag{2.35}$$

This can be seen by considering two cases, depending on whether the minimum is $|x - y_1|$ or $|x_2 - y|$, and noting that $|y_1 - x_2| \leq L_k$. As a consequence, a combinatoric factor

$$c_d L_k^d \min\{|x - y_1|, |x_2 - y|\}^{(d+1)} \leq \tilde{c}_d L_{\min}^{2d+1} \tag{2.36}$$

suffices to control the sums. Here $L_{\min} \equiv \min\{L_a, L_b\}$, $L_{\max} \equiv \max\{L_a, L_b\}$, with $L_a \equiv |x - y_1| \vee L_{k-3}$, $L_b \equiv |x_2 - y| \vee L_{k-3}$. Hence we may bound the $m = 2$ term by

$$\sum_{y_1, x_2} Q_{x, y_1}^{(k-1)} Q_{x_2, y}^{(k-1)} \leq \sup_{y_1, x_2} \tilde{c}_d L_{\min}^{2d+1} Q_{x, y_1}^{(k-1)} Q_{x_2, y}^{(k-1)} \leq \sup_{y_1, x_2} \tilde{c}_d L_{\min}^{2d+1} \tilde{\varepsilon}^{L_a^\chi} \tilde{\varepsilon}^{L_b^\chi}. \tag{2.37}$$

Note that

$$|x - y| \vee L_{k-2} \leq L_a + L_k + L_b = L_{\min} + L_k + L_{\max}, \tag{2.38}$$

since as indicated above, $|x_2 - y_1| \leq L_k$ by connectivity of B. Let us write

$$\tilde{c}_d L_{\min}^{2d+1} \tilde{\varepsilon}^{L_a^\chi} \tilde{\varepsilon}^{L_b^\chi} = \tilde{c}_d L_{\min}^{2d+1} \tilde{\varepsilon}^{L_{\min}^\chi} \tilde{\varepsilon}^{L_{\max}^\chi} \leq \tfrac{1}{4} \tilde{\varepsilon}^{(4/5) L_{\min}^\chi} \tilde{\varepsilon}^{L_{\max}^\chi}, \tag{2.39}$$

which holds for small $\tilde{\varepsilon}$. We claim that

$$\tfrac{4}{5} L_{\min}^\chi + L_{\max}^\chi > (L_{\min} + L_k + L_{\max})^\chi. \tag{2.40}$$

The worst case for this inequality is when L_a, L_b are at their minimum possible values, L_{k-3} (as one can check by differentiating). Then it reduces to the inequality $1.8 > 10^\chi$, or $\chi < \log_{10} 1.8 \approx .255$. Recalling that $\psi = \tfrac{2}{3}$ and $\chi = \psi/3$, we can confirm this inequality, and then (2.38),(2.40) imply a bound

$$\tfrac{1}{4} \tilde{\varepsilon}^{(|x-y| \vee L_{k-2})^\chi} = \tfrac{1}{4} \varepsilon^{q_{k-1}(|x-y| \vee L_{k-2})^\chi} < \tfrac{1}{4} \varepsilon^{q_k(|x-y| \vee L_{k-2})^\chi} \tag{2.41}$$

on the term $m = 2$.

The terms $m > 2$ may be handled inductively, since the bound just proven is sufficient to reproduce the argument when summing each pair y_i, x_{i+1} in turn, and we obtain a bound $2^{-m} \tilde{\varepsilon}^{(|x-y| \vee L_{k-2})^\chi}$ on the mth term.

Now consider the term $m = 1$, *i.e.* the first term in (2.33). This corresponds to a situation where x, y are already contained in the same block on scale $k - 1$. As with the $m \geq 2$ terms, we have a bound given by $Q_{x,y}^{(k-1)}$ on that scale:

$$Q_{x,y}^{(k-1)} \leq \varepsilon^{q_{k-1}(|x-y| \vee L_{k-3})^\chi}. \tag{2.42}$$

We need to improve this by replacing L_{k-3} with L_{k-2}. For this, we can assume that $|x - y| \leq L_{k-2}$.

First, suppose the scale $k-1$ block containing x, y is composed of 2 or more scale $k-2$ blocks. We obtain a bound which is more than sufficient, because we receive two or more factors of $\varepsilon^{q_{k-2}L_{k-4}^\chi}$ from (2.22), which leads to $\varepsilon^{2q_{k-2}L_{k-4}^\chi} < \varepsilon^{q_{k-1}L_{k-2}^\chi}$. (Here we use the fact that $\chi = \frac{2}{9}$, so $2L_{k-4}^\chi = 2^{1-4/9}L_{k-2}^\chi$, and the excess over L_{k-2}^χ will allow us to control the sum over relative positions of the two blocks as in the arguments for (2.35)-(2.41) above.) We should also consider the case $k = 2$, for which we have at least a factor $\varepsilon^{1/5} \le \varepsilon^{q_1 L_0^\chi}$ for a single block on scale 1, as explained in the proof of (2.22) for $k = 1$.

On the other hand, suppose the $(k-1)$-block containing x, y consists of a single $(k-2)$-block. If $k = 3$, then $k - 2 = 1$ and $\varepsilon^{1/5} < \varepsilon^{1/6} = \varepsilon^{q_2 L_1^\chi}$, so the desired bound holds. If not, then at least one of two cases must hold: $n > L_{k-2}^{2\psi/5}$, or $\mathrm{diam}(B) > L_{k-2}$. (Since the block does not join up with other blocks after two jumps in scale $L_{k-2} \to L_k$, its distance from other blocks must be at least $4L_{k-2}$. As we are working on the case where B is *not* isolated on scale $k - 2$, we must violate one of the *other* conditions of Definition 2.1 – provided $k - 2 \ge 2$.) For the case $n > L_{k-2}^{2\psi/5}$, we have by (2.31) a bound

$$\sum_{B_1 \text{ containing } x \text{ and } y} \tilde{P}^{(k-1)}(B_1)\varepsilon^{-q_{k-1}n(B_1)} \exp\left(-|\log\varepsilon|q_{k-1}L_{k-2}^{-\psi/15}n(B_1)\right) \qquad (2.43)$$

on the first term of (2.33). The sum over B_1 is bounded by (2.42), and we pick up an extra factor

$$\varepsilon^{q_{k-1}L_{k-2}^{-\psi/15}n(B_1)} \le \varepsilon^{q_{k-1}L_{k-2}^{\psi/3}} = \varepsilon^{q_{k-1}L_{k-2}^\chi}, \qquad (2.44)$$

which includes the desired improvement $L_{k-3} \to L_{k-2}$. Finally, consider the case of a block on scale $k - 2$ with diameter $> L_{k-2}$. If this is the case, we must be able to find x', y' in the block with $|x' - y'| > L_{k-2}$. Then we can use (2.22), which becomes

$$Q_{x',y'}^{(k-2)} \le \varepsilon^{q_{k-2}|x'-y'|^\chi} \le \varepsilon^{q_{k-2}L_{k-2}^\chi}, \qquad (2.45)$$

and after summing over $\le (2L_{k-2}+1)^d$ possibilities for x' and for y', given that x, y are in the block, we obtain a bound $\varepsilon^{q_{k-1}L_{k-2}^\chi}$, as in the other cases. (The drop $q_{k-2} \to q_{k-1}$ allows us to use the additional factor $\varepsilon^{q_{k-2}L_{k-3}^{-\psi/20}L_{k-2}^\chi}$ to control the sums – see (2.46) below.) Altogether, we have considered up to three cases for any particular k. Let us allow $|x - y| > L_{k-2}$ again. We obtain the following bound on the $m = 1$ term:

$$3\varepsilon^{q_{k-1}(|x-y|\vee L_{k-2})^\chi} \le 3\varepsilon^{q_k(|x-y|\vee L_{k-2})^\chi}\varepsilon^{q_{k-1}L_{k-2}^{-\psi/20}L_{k-2}^\chi} < \tfrac{1}{4}\varepsilon^{q_k(|x-y|\vee L_{k-2})^\chi}. \qquad (2.46)$$

We return to a consideration of $Q_{x,y}^{(k),2}$, which consists of terms from (2.21) that are isolated on scale $k - 2$, which means that $k \ge 4$. In this case we have from (2.20) that $\tilde{P}^{(k)}(B) = \varepsilon^{L_k^{3\psi/5}/3}$. We need to sum over all possible blocks of diameter $\le L_{k-2}$ and weight by $\varepsilon^{-q_k n}$ to obtain $Q_{x,y}^{(k),2}$. We obtain a bound

$$\varepsilon^{L_{k-2}^{3\psi/5}/3}(2L_{k-1}+1)^{dn}\varepsilon^{-q_k n} < \varepsilon^{L_{k-2}^{3\psi/5}/4} < \tfrac{1}{4}\varepsilon^{q_k L_{k-2}^\chi} = \tfrac{1}{4}\varepsilon^{q_k(|x-y|\vee L_{k-2})^\chi}, \qquad (2.47)$$

using $dn\log(2L_{k-2}+1) < c_d n L_{k-2}^{\psi/5} \le c_d L_{k-2}^{3\psi/5}$ and $q_k n \le q_k L_{k-2}^{2\psi/5} < L_{k-2}^{3\psi/5}/12$ (and $q_k < \frac{1}{12}$ for $k \ge 4$, which can readily be checked from the recursion). Note that we are

using the Proposition 2.3 bound to control the sum over all substructure of B. This is because Proposition 2.3 cannot be used simultaneously for a block and for its subblocks, due to lack of independence. However, the bound for the whole block is much better than what had been obtained earlier for subblocks, and (2.47) demonstrates that it is adequate for our purposes here.

To complete the proof of the theorem, we gather the bounds (2.46), (2.47) and the bound $2^{-m}\varepsilon^{q_k(|x-y|\vee L_{k-2})^\chi}$ proven above on the mth term, $m \geq 2$. The desired bound (2.22) follows, since $\frac{1}{4} + \frac{1}{4} + \sum_{m\geq 2} 2^{-m} = 1$. $\qquad\square$

We may now estimate $\mathcal{P}_{x,y}^{(k)}$, the probability that x and y belong to the same resonant block on scale k.

Corollary 2.6. *Given $\psi = \frac{2}{3}$, $\chi = \psi/3$, let ε be sufficiently small. Then there is a $q > 0$ such that*

$$\mathcal{P}_{x,y}^{(k)} \leq \varepsilon^{q(|x-y|\vee L_{k-2})^\chi}. \tag{2.48}$$

Proof. By construction, the probabilities $(3\sqrt{\varepsilon})^{n(B)}$ and $\varepsilon^{L_k^{3\psi/5}/3}$ in (2.20) correspond to independent events, because whenever a block is resonant on a scale k, its probability bound from Proposition 2.3 replaces all previous probability factors for its subblocks. So any two subblocks with a factor $\varepsilon^{L_k^{3\psi/5}/3}$ appearing in (2.20) are separated, ensuring independence. (Independence is a result of the truncation of the random walk expansion to form $\tilde{F}_E^{(k-1)}(B)$, including only graphs that remain within a distance $< L_{k-2}$ of B, so that there is no overlap between sets of variables $\mathbf{u}$ on which each resonance event depends.) At each length scale, isolated blocks must be resonant (otherwise they would be removed before forming blocks at the next scale) and hence can be associated with a probability factor $\varepsilon^{L_k^{3\psi/5}/3}$ from Proposition 2.3. Thus all the probability factors in (2.20) correspond to independent events that must hold if B is a component of $R^{(k)}$. Hence $\tilde{P}^{(k)}(B)$ is a bound on the probability that B is a component of $R^{(k)}$.

Since the weights in (2.21) are bounded below by 1, we see that (2.22) becomes a bound on the sum of $\tilde{P}^{(k)}(B)$ over B containing x, y. The corollary follows, since $q_k > q > 0$. $\qquad\square$

The underlying mechanisms at work in the proof of Theorem 2.4 are as follows. First, since we are only seeking fractional exponential decay, the procedure can tolerate giving up some fraction of the decay distance with each new length scale (which happens because of the gaps that were required between resonant components). Smallness of the probability of a component comes in one of two ways. If a component is isolated, with small volume and diameter, the basic proposition on resonance probability can be used to bring in new smallness to control its greater positional freedom – it gives an estimate that decays as a fractional exponential of the length scale. On the other hand, if the component's volume or diameter is large, or it is not too far from other components, the decay can be obtained from previous scales. We need to be careful not to give up too much volume decay for this purpose in each step, so that we can obtain bounds that are uniform in the step index k.

2.4 Random Walk Expansion

The random walk expansion is based on the Schur complement at level k. We put

$$F_\lambda^{(k-1)} = \begin{pmatrix} A^{(k)} & B^{(k)} \\ C^{(k)} & D^{(k)} \end{pmatrix}, \qquad (2.49)$$

where the blocks are determined by the decomposition of $R^{(k-1)}$ into $R^{(k)}$ (upper-left block) and $R^{(k-1)} \setminus R^{(k)}$ (lower-right block). We may now define the kth Schur complement for $|\lambda - E| \le \varepsilon_k/2$:

$$F_\lambda^{(k)} = A^{(k)} - B^{(k)}(D^{(k)} - \lambda)^{-1}C^{(k)}, \qquad (2.50)$$

Let us write

$$D^{(k)} = W^{(k)} + V^{(k)}, \qquad (2.51)$$

where $W^{(k)}$ is block diagonal, each block being $\tilde{F}_\lambda^{(k-1)}(B)$ for some B. This means that $V^{(k)}$ consists of the long graphs not included in $\tilde{F}_\lambda^{(k-1)}(B)$, producing matrix elements both within blocks and between blocks.

We show below in Theorem 2.9 that

$$\|\tilde{F}_\lambda^{(k-1)}(B) - \tilde{F}_E^{(k-1)}(B)\| \le c_d \frac{\gamma}{\varepsilon}|\lambda - E|, \qquad (2.52)$$

which is less than $\varepsilon_k/6$, because $|\lambda - E| \le \varepsilon_k/2$. Since all the blocks of $R^{(k-1)} \setminus R^{(k)}$ are non-resonant,

$$\text{dist}\big(\text{spec}\,\tilde{F}_E^{(k-1)}(B), E\big) \ge \varepsilon_k \equiv \varepsilon^{L_k^\psi}, \qquad (2.53)$$

and so

$$\|(W^{(k)} - \lambda)^{-1}\| \le 3\varepsilon_k^{-1}. \qquad (2.54)$$

Hence, as in the first step, we may expand $(D^{(k)} - \lambda)^{-1}$ in a Neumann series:

$$[(D^{(k)} - \lambda)^{-1}]_{xy} = \sum_{g_k:x \to y} \prod_{i=1}^m [(W^{(k)} - \lambda)^{-1}]_{x_i \tilde{x}_i} \prod_{j=1}^{m-1} V_{\tilde{x}_j, x_{j+1}}^{(k)}. \qquad (2.55)$$

Here $g_k = \{x = x_1, \tilde{x}_1, x_2, \tilde{x}_2, \ldots, x_m, \tilde{x}_m = y\}$, with each $x_i, \tilde{x}_i$ in the same block. Note that $V_{x,y}^{(k)}$ is given by a sum of graphs contributing to $F_\lambda^{(k-1)} - \oplus_B \tilde{F}_\lambda^{(k-1)}(B)$, where

$$F_\lambda^{(k-1)} = A^{(k-1)} - B^{(k-1)}(D^{(k-1)} - \lambda)^{-1}C^{(k-1)}. \qquad (2.56)$$

Thus, when it is useful, we may unwrap any matrix element into a sum of multigraphs, since each step at any level consists of a sum of graphs on the previous level. Thus, the matrices $A^{(k-1)}, B^{(k-1)}, C^{(k-1)}, D^{(k-1)}$ in (2.56) are themselves blocks of $F_\lambda^{(k-2)}$, which is given by $A^{(k-2)} - B^{(k-2)}(D^{(k-2)} - \lambda)^{-1}C^{(k-2)}$. We may continue down to graphs at level 1.

We will also prove bounds on the matrices that generate the eigenfunctions. Recall from the fundamental lemma on the Schur complement that if $\varphi^{(k)}$ is an eigenvector of $F_\lambda^{(k)}$ with eigenvalue λ, then

$$\varphi^{(k-1)} = \left(\varphi^{(k)}, -(D^{(k)} - \lambda)^{-1} C^{(k)} \varphi^{(k)}\right) \tag{2.57}$$

is an eigenvector of $F_\lambda^{(k-1)}$ with the same eigenvalue. This process may be repeated to extend the eigenvector $\varphi^{(k)}$ all the way out to the original lattice Λ^*, that is, to produce $\varphi^{(0)}$, an eigenvector of H. Let us write

$$\varphi^{(0)} = G_\lambda^{(k)} \varphi^{(k)}, \tag{2.58}$$

and then we may give a multigraph expansion for $G_\lambda^{(k)}$ in the same manner as was just described for $F_\lambda^{(k)}$. Indeed, the same operators $C^{(k)}, D^{(k)}$ appear when unwrapping (2.57). The matrix $G_\lambda^{(k)}$ can be thought of as half of $F_\lambda^{(k)}$ since it descends to the original lattice, but does not climb back up to $R^{(k)}$.

One may visualize the multigraph expansion on the original lattice as an ordinary random walk of nearest-neighbor steps, except that when it enters a block on level k, there is a matrix element of $(W^{(k)} - \lambda)^{-1}$ that produces an intra-block jump. This structure is present already at the first level, see (2.8); but there we had only single-position blocks. The larger blocks that appear in later steps create gaps in decay, since there is no decay from x_i to $\tilde{x}_i$ in the following estimate for a walk passing through a block B of $R^{(k-1)} \setminus R^{(k)}$:

$$\sum_{\tilde{x}_i} \left|[(W^{(k)} - \lambda)^{-1}]_{x_i \tilde{x}_i}\right| \leq n(B) \cdot 3\varepsilon_k^{-1} \leq 3 L_k^{2\psi/5} \varepsilon^{-L_k^\psi}. \tag{2.59}$$

Here we use (2.54), and note that there are just $n(B)$ choices for $\tilde{x}_i$, given x_i. As the block is isolated on scale k, its volume is bounded by $L_k^{2\psi/5}$. The main issue with controlling the multigraph expansion is making sure that the gaps in decay and the large factors from the bound (2.59) do not spoil too much the exponential decay proven in step 1. Our constructions ensure that there are gaps of size $4L_k$ between blocks at scale k. Hence, whatever rate of decay is proven at scale $k-1$, some small fraction will be lost at scale k. This is why we are led to fractional exponential decay estimates. Keep in mind that as far as the step k random walk expansions are concerned, the configuration of blocks up to that level $(R^{(1)}, R^{(2)}, \ldots, R^{(k)})$ is fixed.

We now state our main theorem on graphical bounds. Let $\mathcal{S}_{x,z,y}^{(k)}$ denote the sum of the absolute values of all multigraphs for $B^{(k)}(D^{(k)} - \lambda)^{-1} C^{(k)}$ that go from x to y and that contain z. Here, x, y are positions in $R^{(k)}$, and z is a position in $R^{(k)} \setminus R^{(k-1)}$. We say that a multigraph contains z if any of the blocks that it passes through contain z.

Theorem 2.7. *With $\psi = \frac{2}{3}, \phi = \frac{1}{4}$, and $\varepsilon = \gamma^\phi$ sufficiently small, assume that $|\lambda - E| \leq \varepsilon_k/2$. Then*

$$\mathcal{S}_{x,z,y}^{(k)} \leq \gamma^{[(|x-z|+|z-y|)\vee L_k]^\psi} \cdot 2^{-k} \tag{2.60}$$

$$\sum_{j \leq k} \mathcal{S}_{x,z,y}^{(j)} \leq \gamma^{[(|x-z|+|z-y|)\vee 1]^\psi}. \tag{2.61}$$

20

Proof. For the case $k = 1$ we obtain

$$\mathcal{S}_{x,z,y}^{(k)} \leq \left(\frac{c_d\gamma}{\varepsilon}\right)^{(|x-z|+|z-y|)\vee 2} \cdot \varepsilon, \tag{2.62}$$

as in the proof of (2.10). Let $L = (|x-z|+|z-y|)\vee 1$, then one has a power $\frac{3}{4}(L\vee 2)+\frac{1}{4}$ of γ in (2.62). This is greater than the required power $L^{2/3}$ for (2.60), as one can readily verify, checking the first few values of L by hand. Hence for γ sufficiently small, we obtain (2.60).

Let us now assume (2.60) for $k - 1$. Using (2.54), one may expand $B^{(k)}(D^{(k)} - \lambda)^{-1}C^{(k)}$ to obtain a graph connecting x to y traversing m intermediate blocks. In the figure, each block has a matrix element of $(W^{(k)} - \lambda)^{-1}$ as in (2.59). The extra point z must lie in one of the sections of the graph (as in Figure 2) or in one of the blocks. Because the blocks are isolated, and x, y are in $R^{(k)}$, each section of the graph traverses
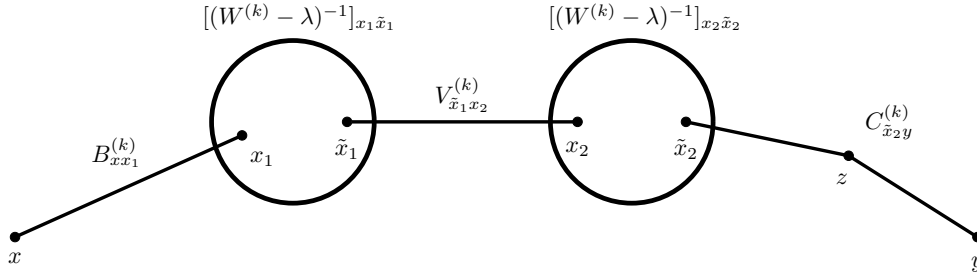

Figure 2: A graph from x to y in the step k random walk expansion

a distance $\geq 4L_k$, and the blocks themselves have diameter $\leq L_k$. We sum over the pairs $(x_i, \tilde{x}_i)$ in turn in a manner similar to what was done for the probability estimate. In Figure 2, for example, the sum over x_i may be controlled with a combinatoric factor $c_d|x-x_i|^{d+1}$ and (2.59) controls the sum over $\tilde{x}_1$ and results in a factor $3L_k^{2\psi/5}\varepsilon^{-L_k^{\psi}}$. Applying the inductive hypothesis to $B^{(k)}$, $C^{(k)}$, and $V^{(k)}$ (which come from $F_\lambda^{(k-1)}$), we obtain a bound on the portion from x to x_2:

$$\gamma^{|x-x_1|^\psi} c_d|x-x_1|^{d+1} \cdot 3L_k^{2\psi/5}\varepsilon^{-L_k^{\psi}}\gamma^{|\tilde{x}_1-x_2|^\psi}, \tag{2.63}$$

noting that $|x - x_1| \equiv L_a \geq 4L_k$, and $|\tilde{x}_1 - x_2| \equiv L_b \geq 4L_k$. We claim that

$$L_a^\psi - \phi L_k^\psi + L_b^\psi > (L_a + 2L_k + L_b)^\psi \geq (|x - x_2| \vee 4L_k)^\psi. \tag{2.64}$$

The first inequality may be reduced to the minimal case $L_a = L_b = 4L_k$ by comparing derivatives of both sides. Then it becomes $4^\psi - \phi + 4^\psi > 10^\psi$, which is easy to verify for $\psi = \frac{2}{3}, \phi = \frac{1}{4}$. The second inequality is clear. Since the inequality is strict in (2.64), there is a sufficient leftover power of γ to control the factors of $|x - x_1|$ and L_k and obtain an overall bound

$$\gamma^{(|x-x_2|\vee 4L_k)^\psi} \cdot 4^{-k}, \tag{2.65}$$

which is sufficient to repeat the argument when summing over $(x_2, \tilde{x}_2)$ in the next block. In this manner we may proceed sequentially through a chain of m blocks.

The point z must lie either in one of the blocks or in one of the connecting graphs. If a connecting graph is required to pass through z, then we may take, for example, $L_b = |\tilde{x}_1 - z| + |z - x_2|$ and the argument goes through, achieving decay along a path that includes z. If z is in a block, then the path from x to z to y may entail a double traverse of the block, but this is covered by the factor of 2 in (2.64). Furthermore, there is also the possibility of a self-line, where $V^{(k)}$ connects a block to itself. Such graphs contain a point z' that is a distance $\geq L_{k-1}$ from B, so $|x - z'| + |z' - y| \geq 2L_{k-1} = L_k$, and then summing (2.60) over z' and y gives a bound $c_d L_k^{d+1} L_k^d \gamma^{L_k^\psi}$ on the sum of all such graphs. Thus we may account for the self-lines with a factor 1.1, say, per block. Another factor per block controls the sum over which part of the graph contains z. These considerations allow us to obtain from (2.64) the desired decay distance $(|x - y| + |y - z|) \vee L_k$, along with a factor 3^{-km} for a chain of m blocks from x to y. Summing this over $m \geq 1$ leads to the desired bound, (2.60). The bound (2.61) on the sum of graphs on all scales follows immediately. This completes the proof. $\qquad\square$

Corollary 2.8. *Under the same assumptions as Theorem 2.7,*

$$\|F_\lambda^{(k)} - \oplus_B \tilde{F}_\lambda^{(k)}(B)\| \leq (c_d \gamma)^{L_k^\psi}. \tag{2.66}$$

Proof. Graphs contributing to the difference go from x to y via a point z such that $|x - z| \geq L_{k-1}$, $|y - z| \geq L_{k-1}$. We may bound the norm by estimating the maximum absolute row sum of the matrix. This means fixing x and taking the sum over z and y of (2.61). A combinatoric factor $c_d L^{2d+2}$ controls the sums, where $L = |x - z| + |y - z| \geq 2L_{k-1} = L_k$, and (2.66) follows. $\qquad\square$

We also need to control the difference $\tilde{F}_\lambda^{(k)}(B) - \tilde{F}_E^{(k)}(B)$ in norm, so that when isolated blocks are defined via the condition $\text{dist}\big(\text{spec}\,\tilde{F}_E^{(k-1)}(B), E\big) > \varepsilon_k$, it is still safe to build the random walk expansion for the Schur complement with respect to λ.

Theorem 2.9. *With $\psi = \frac{2}{3}, \phi = \frac{1}{4}$, and $\varepsilon = \gamma^\phi$ sufficiently small, assume that $|\lambda - E| \leq \varepsilon_k/2$. Then*

$$\|\tilde{F}_\lambda^{(k)}(B) - \tilde{F}_E^{(k)}(B)\| \leq c_d \frac{\gamma}{\varepsilon} |\lambda - E|. \tag{2.67}$$

Proof. We have $F_\lambda^{(k)} = A^{(k)} - B^{(k)}(D^{(k)} - \lambda)^{-1} C^{(k)}$. In addition to the explicit appearance of λ, the matrices $A^{(k)}, B^{(k)}, C^{(k)}, D^{(k)}$ depend on λ for $k \geq 2$. We already have control of the graphs contributing to these expressions by Theorem 2.7. An inductive argument will allow us to control also the difference when we change λ to E.

We begin by proving an analog of Theorem 2.7 to control the sum of differences of graphs, *i.e.* each graph is evaluated at λ and at E and the difference taken. Let $\tilde{S}_{x,y}^{(k)}$ denote the sum of the absolute values of all difference multigraphs that contribute to $[B^{(k)}(D^{(k)} - \lambda)^{-1} C^{(k)}]_{xy}$. We claim that

$$\tilde{S}_{x,y}^{(k)} \leq \frac{1}{\varepsilon_k} \gamma^{(|x-y| \vee L_k)^\psi} \cdot 2^{-k} |\lambda - E|, \tag{2.68}$$

22

and hence that

$$\sum_{j \leq k} \tilde{\mathcal{S}}_{x,y}^{(j)} \leq \frac{1}{\varepsilon} \gamma^{(|x-y| \vee 1)^{\psi}} |\lambda - E|. \tag{2.69}$$

Consider the case $k = 1$. Redoing the proof of (2.10) for differences, we obtain a sum of graphs wherein a difference

$$[(W^{(1)} - \lambda)^{-1} - (W^{(1)} - E)^{-1}]_{x_i \tilde{x}_i} = (\lambda - E)[(W^{(1)} - \lambda)^{-1}(W^{(1)} - E)^{-1}]_{x_i \tilde{x}_i} \tag{2.70}$$

appears in place of the corresponding matrix element of $(W^{(1)} - \lambda)^{-1}$ or $(W^{(1)} - E)^{-1}$. Compared to what we had before, there is a combinatoric factor of m for the choice of i, a factor $|\lambda - E|$, and an extra $1/\varepsilon$ from the additional $(W^{(1)} - E)^{-1}$. We obtain in place of (2.62)

$$\tilde{\mathcal{S}}_{x,y}^{(1)} \leq \left(\frac{c_d \gamma}{\varepsilon}\right)^{(|x-y| \vee 2)} |\lambda - E| \leq \frac{1}{2\varepsilon} \gamma^{(|x-y| \vee 1)^{\psi}} |\lambda - E|, \tag{2.71}$$

which verifies (2.68). For the second inequality of (2.71), let $L = |x - y| \vee 1$. One can check that $\frac{3}{4}(L \vee 2) > L^{2/3} - \frac{1}{4}$, and then the claim may be verified by counting powers of γ.

For step k, we apply the difference operation to each term of (2.55), and also to $B^{(k)}, C^{(k)}$ in the expression $B^{(k)}(D^{(k)} - \lambda)^{-1}C^{(k)}$. Each matrix $W^{(k)}, V^{(k)}, B^{(k)}, C^{(k)}$ is covered by (2.69), by induction, and this leads to an incremental factor of $\varepsilon^{-1}|\lambda - E|$, compared to before. When we difference the explicit factors of λ in (2.55), we obtain as in (2.70) a new factor of $(W^{(k)} - E)^{-1}|\lambda - E|$. This leads to an incremental factor $\varepsilon_k^{-1}|\lambda - E|$, compared to before, coming from the bound $\|(W^{(k)} - E)^{-1}\| \leq \varepsilon_k^{-1}$. Thus in all cases, we get no worse than a factor $\varepsilon_k^{-1}|\lambda - E|$. This completes the proof of (2.68), (2.69).

Note that (2.69) provides an estimate on the matrix elements of $F_\lambda^{(k)} - F_E^{(k)}$, so that

$$\sum_y \left| [F_\lambda^{(k)} - F_E^{(k)}]_{xy} \right| \leq c_d \frac{\gamma}{\varepsilon} |\lambda - E|, \tag{2.72}$$

and hence

$$\|F_\lambda^{(k)} - F_E^{(k)}\| \leq c_d \frac{\gamma}{\varepsilon} |\lambda - E|, \tag{2.73}$$

The same bound applies to $\|\tilde{F}_\lambda^{(k)}(B) - \tilde{F}_E^{(k)}(B)\|$, since in this case we are just looking at a subset of the collection of multigraphs (the ones that remain within a distance $< L_{k-1}$ of B). $\qquad\square$

In order to prove decay of eigenfunctions, we need a bound similar to Theorem 2.7 on the sum of multigraph contributing to $G_\lambda^{(k)}$, the eigenfunction-generating kernel. Let $\mathcal{G}_{x,y}^{(k)}$ denote the sum of the absolute values of all multigraphs for $G_\lambda^{(k)}$ – see (2.57),(2.58) for its definition.

23

Theorem 2.10. *With $\psi = \frac{2}{3}, \phi = \frac{1}{4}$, and $\varepsilon = \gamma^\phi$ sufficiently small, assume that $|\lambda - E| \leq \varepsilon_k/2$. Then*

$$\mathcal{G}_{x,y}^{(k)} \leq \left(\frac{\gamma}{\varepsilon}\right)^{(|x-y|\vee L_k)^\psi} \cdot 2^{-k} \tag{2.74}$$

$$\sum_{j \leq k} \mathcal{G}_{x,y}^{(j)} \leq \left(\frac{\gamma}{\varepsilon}\right)^{(|x-y|\vee 1)^\psi}. \tag{2.75}$$

Note that along with the graphical terms controlled by this theorem, $G_\lambda^{(k)}$ contains an identity matrix at the block for $R^{(k)}$, see (2.57). The theorem only controls the nontrivial graphs in $G_\lambda^{(k)}$.

Proof. We have already established in Theorem 2.7 the needed bounds on $D^{(k)}, C^{(k)}$, and also $W^{(k)}, V^{(k)}$, which appear in the random walk expansion (2.55) for $(D^{(k)} - \lambda)^{-1}$. When estimating $(D^{(k)} - \lambda)^{-1}C^{(k)}$, we have a picture similar to Figure 2, except without the $B^{(k)}$ line. In the proof of Theorem 2.7, gaps in decay and factors ε_k^{-1} at blocks were controlled by (2.64). There we had decay on either side of each block, which allowed us to prove uniform fractional exponential decay. Here we need to work with just the line on one side. Consider the situation in Figure 3. Allowing for combinatoric factors
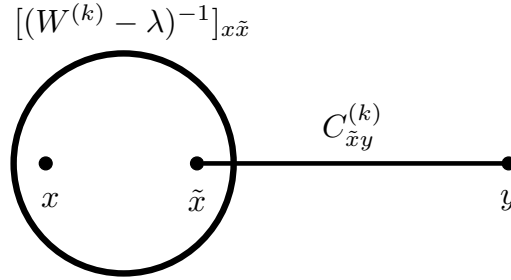

Figure 3: A graph from the eigenfunction-generating kernel.

as in (2.63), we obtain a bound

$$3L_k^{\psi/2}\varepsilon^{-L_k^\psi}\gamma^{|\tilde{x}-y|^\psi}, \tag{2.76}$$

with $|\tilde{x} - y| \equiv L_a \geq 4L_k$. We claim that

$$L_a^\psi - \phi L_k^\psi > \tfrac{3}{4}(L_a + L_k)^\psi. \tag{2.77}$$

The worst case for this inequality is when $L_a = 4L_k$, where it reduces to $4^\psi - \phi > \frac{3}{4}5^\psi$, which can easily be verified for $\phi = \frac{1}{4}, \psi = \frac{2}{3}$. Hence we may bound (2.76) by

$$\left(\frac{\gamma}{\varepsilon}\right)^{(|x-y|\vee L_k)^\psi} \cdot 4^{-k}. \tag{2.78}$$

Repeating this argument along a chain of blocks, and then along the chain of operators $\prod_j(-(D^{(j)} - \lambda)^{-1}C^{(j)})$ that go into the definition of $G_\lambda^{(k)}$, we obtain (2.74). The bound (2.75) on the sum over scales follows immediately. $\qquad\square$

24

3 Proofs of Main Results

We have established probabilistic estimates on resonant blocks, giving decay both in the spatial extent and in the volume of the block. Furthermore, we have obtained graphical estimates that imply decay of eigenfunctions away from resonant blocks. Let us put these tools to work by proving our main theorems. In the first section we prove our theorem on density of states. In Section 3.2, we introduce an "energy-following procedure" that serves to construct every eigenstate through a process of successive approximation. In Section 3.3, we prove the needed probabilistic bound on discriminants. In Section 3.4, we complete the proof of fractional exponential decay of the eigenfunction correlator. In Section 3.5, we extend the method to complete the level-spacing result.

3.1 Density of States

Theorem 3.1. *There exists a constant $b > 0$ such that for γ sufficiently small the following is true. For any $0 < \delta \le \gamma^{1/4}$, any rectangle Λ, and any interval $I = [E - \delta/2, E + \delta/2]$, let $N(I)$ denote the number of eigenvalues of H in I. Then*

$$\mathbb{E}\, N(I) \le 4|\Lambda| \exp(-b|\log\gamma|^{2/3}|\log\delta|^{1/3}). \tag{3.1}$$

The focus of this theorem is on the behavior as $\delta \to 0$, in particular we see that $\mathbb{E}N(I)$ vanishes faster than any power of $1/|\log\delta|$. If one considers instead a "macroscopic" interval with $\delta > \varepsilon = \gamma^\phi = \gamma^{1/4}$, then the following bound holds:

$$\mathbb{E}\, N(I) \le 4\sqrt{\delta}|\Lambda|. \tag{3.2}$$

This estimate may be obtained by a small modification of step 1. Define $R^{(1)}$ using the criterion that $t_x \in [E - \delta, E + \delta]$. Then as in (2.2) we have a bound of $2\sqrt{\delta}$ on the probability that a position is in $R^{(1)}$. Weyl's inequality limits the movement of eigenvalues to $O(\gamma)$, hence $N(I)$ is bounded by $2|R^{(1)}|$. The expectation of $|R^{(1)}|$ is the sum on x of the probability that $x \in R^{(1)}$, so (3.2) follows from (2.2).

Proof of Theorem 3.1. We follow the argument just given for (3.2), but work with $R^{(k)}$ instead of $R^{(1)}$. Here k is defined by the condition $\varepsilon_{k+1} < \delta \le \varepsilon_k$ (recall that $\varepsilon_k \equiv \varepsilon^{L_k^\psi}$). By assumption, $\delta \le \gamma^{1/4} = \varepsilon = \varepsilon_1$. The probability that $x \in R^{(k)}$ is bounded by

$$\mathcal{P}_{x,x}^{(k)} \le \varepsilon^{qL_{k-2}^\chi} \le \varepsilon^{\tilde{q}2^{k\chi}} = \varepsilon^{\tilde{q}(2^{k\psi})^{1/3}}, \tag{3.3}$$

using (2.48), $\tilde{q} \equiv q/8^\chi$, $\chi = \psi/3$. We may write this as

$$\exp(-\tilde{q}|\log\varepsilon||\log_\varepsilon \varepsilon_{k+1}|^{1/3}) \le \exp(-\tilde{q}|\log\varepsilon||\log_\varepsilon \delta|^{1/3})$$
$$= \exp(-\tilde{q}|\log\varepsilon|^{2/3}|\log\delta|^{1/3}), \tag{3.4}$$

and after summing over x, we obtain (3.1). $\square$

Remark 1. The bounds (3.1),(3.2) imply the following unified estimate, valid for $0 < \delta \le 1$:

$$\mathbb{E}\, N(I) \le 4|\Lambda| \exp(-b|\log(\gamma \vee \delta)|^{2/3}|\log\delta|^{1/3}). \tag{3.5}$$

Remark 2. The fractional power of $|\log \delta|$ inside the exponential is a result of the discrepancy between χ (the probability exponent) and ψ (the graphical exponent). The relation $\chi = \psi/3$ (or something similar) is necessitated ultimately by the $\frac{1}{2n}$ exponent in (2.18), which leads to the volume limitation in Definition 2.1. This leads to limits on χ from (2.19) and (2.44), depending on how large n is.

3.2 Energy-Following Procedure

We will need a modified procedure that chooses energy windows to contain successive approximations to eigenvalues of H. Observe that the spectrum of H is necessarily contained in the union of $2d\gamma$-neighborhoods of each $\pm t_x, x \in \Lambda$. (Note that $\|V\| = \|J\| \le 2d\gamma$, see (1.2),(1.3).) Let us choose one particular x and put E_1 equal to one of the two eigenvalues $\pm t_x$ of h_x. Obviously E_1 depends on x and on the choice of eigenvalue, but we suppress the dependence in the notation. We may then define the Schur complement $F_{E_1}^{(1)}$, as described in Section 2.1.

Recall that blocks were formed by taking connected clusters of sites such that $\mathrm{dist}(E_1, \{t_y, -t_y\}) \le \varepsilon$ (based on nearest-neighbor connections). Note that if we wish to estimate the probability that x, y are contained in the same block, as before there must be a sequence of distinct nearest-neighbor positions in $R^{(1)}$ connecting x to y. The probability that any of these positions (other than x) is in $R^{(1)}$ is bounded by $3\sqrt{\varepsilon}$, as before. By construction, x is in $R^{(1)}$, so there is no small probability associated with the event $\{x \in R^{(1)}\}$. Thus we need to define $\tilde{\mathcal{P}}_{x,y}^{(1)}$, the probability that y belongs to the block based at x (with a particular choice of E_1 taken as fixed). We have that

$$\tilde{\mathcal{P}}_{x,y}^{(1)} \le (12d\sqrt{\varepsilon})^{|x-y|}. \tag{3.6}$$

Compared with (2.3), we have one fewer factor of $12d\sqrt{\varepsilon}$, for the reason just described. Graphical bounds such as (2.10) are unaffected by the new procedure.

To go on with the second step, we will need the truncated version of the Schur complement matrix. For $|\lambda - E_1| \le \varepsilon/3$ we define as before $\tilde{F}_\lambda^{(1)}(B)$ by restricting the sum of graphs that defines $F_\lambda^{(1)}$, including only those that remain within a distance $< L_0 = \frac{1}{2}$ of B. Thus in this step, we consider no graphs of $B^{(1)}(D^{(1)} - \lambda)^{-1}C^{(1)}$. This means that $\tilde{F}_\lambda^{(1)}(B)$ does not actually depend on λ. But note that interaction terms in V that hop between positions of B are included in $\tilde{F}_\lambda^{(1)}(B)$. Since $\|V\| \le 2d\gamma$, the eigenvalues of $\tilde{F}_\lambda^{(1)}(B)$ are within $2d\gamma$ of the "bare" eigenvalues $\pm t_y, y \in B$. In a manner similar to what was done in the first step, we put E_2 equal to one of the eigenvalues of $\tilde{F}_{E_1}^{(1)}(B)$ in $[E_1 - \varepsilon/3, E_1 + \varepsilon/3]$. There can be no more than $n(B)$ choices for E_2 given our initial choices of (x, E_1). We will discuss the counting of choices $E_1, E_2, \ldots$ below; for now let us observe simply that in view of (2.20), there is a rapid decay of probability with $n(B)$, so the choice of E_2 is under control. (There is a potential for redundancy here, because an eigenvalue may be reached via multiple initial choices of (x, E_1); this is not a problem as long as every eigenvalue is covered at least once.)

The process continues in the kth step in a similar fashion. At this point, there is a sequence of choices $x, E_1, \ldots, E_{k-1}$, and the associated increasing sequence of blocks

containing x, which may be denoted $B_{x,1}, \ldots, B_{x,k-1}$. We look at the spectrum of $\tilde{F}_{\lambda}^{(k-1)}(B_{x,k-1})$ in $[E_{k-1} - \varepsilon_{k-1}/3, E_{k-1} + \varepsilon_{k-1}/3]$, for λ in this same interval. We look for solutions of the equation $\lambda \in \operatorname{spec} \tilde{F}_{\lambda}^{(k-1)}(B_{x,k-1})$, as this closely approximates the exact equation $\lambda \in \operatorname{spec} F_{\lambda}^{(k-1)}$ that determines eigenvalues of H – see the fundamental lemma (we discuss the existence of such solutions below). Theorem 2.9 shows that the matrix $\tilde{F}_{\lambda}^{(k-1)}(B_{x,k-1})$ depends only weakly on λ: it obeys a Lipschitz condition with constant $c_d \gamma / \varepsilon$. By Weyl's inequality, the same is true of its eigenvalues. Hence one may sweep λ through the interval above, and all the solutions to the equation $\lambda \in \operatorname{spec} \tilde{F}_{\lambda}^{(k-1)}(B_{x,k-1})$ in $[E_{k-1} - \varepsilon_{k-1}/3, E_{k-1} + \varepsilon_{k-1}/3]$ may be used as the next set of approximate eigenvalues. We put E_k equal to one such solution. (Note that by requiring $|E_k - E_{k-1}| \leq \varepsilon_{k-1}/3$, we have that $|E_k - E_j| < \varepsilon_j/2$ for $j < k$, because the sum of shifts $\varepsilon_i/3$ for $j \leq i < k$ is less than $\varepsilon_j/2$. This ensures that we never leave the "safe" zone covered by Theorems 2.7, 2.9, and 2.10.) Each choice may then serve as the central energy for the next Schur complement $F_{E_k}^{(k)}$ as defined in Section 2.4. It is important to note that when we shift to E_k, we shift $F_{E_{k-1}}^{(j)} \to F_{E_k}^{(j)}$ for the random walk expansions at level $j < k$ as well. We have the flexibility to do this because in step j, Theorem 2.7 controls graphical expansions constructed at any λ such that $|\lambda - E_j| \leq \varepsilon_j/2$.

There are some new aspects to the probability estimates for this "energy-following" procedure, as compared to the previous fixed-energy procedure. We need to adjust for the fact that by picking E_k close to one of the solutions to $\lambda \in \operatorname{spec} \tilde{F}_{\lambda}^{(k-1)}(B_{x,k-1})$, we lose the ability to obtain smallness of the probability of $B_{x,k-1}$, as $B_{x,k-1}$ is resonant to E_k by construction. More precisely, there is no *new* smallness in the probability estimate; when $B_{x,k-1}$ was formed in the previous step, any nontrivial structure was associated with smallness of probability (as we saw already in (3.6) in step 1). Note, however, that once $B_{x,k-1}$ (the block containing x) is fixed, resonance probabilities for the *other* blocks of $R^{(k-1)}$ involve random variables that are independent of those involved in defining $E_1, \ldots, E_k$ and $B_{x,k-1}$. This is why we always use the truncated matrices $\tilde{F}_{\lambda}^{(j)}(B_{x,j})$ to define the next set of central energies. Due to this independence, the probability estimates for the blocks not containing x (in particular Proposition 2.3) are no different from before. Thus the Theorem 2.4 bound (2.22) holds for blocks not containing x.

We prove a modified version of Theorem 2.4 that applies to the block containing x, for fixed choices of $E_1, \ldots, E_k$. Let $B_{x,k}$ denote the block of x in step k, with $B_{x,0} \equiv \{x\}$. We introduce an altered version of the inductive definition (2.20) that is suited to the energy-following procedure:

$$\hat{P}^{(1)}(B) \equiv \begin{cases} (3\sqrt{\varepsilon})^{n(B)-1}, & \text{if } x \in B; \\ (3\sqrt{\varepsilon})^{n(B)}, & \text{if } x \notin B \end{cases}$$

$$\hat{P}^{(k)}(B) \equiv \begin{cases} \varepsilon^{L_{k-2}^{3\psi/5}/3}, & \text{if } B \text{ is isolated on scale } k-2 \text{ and } x \notin B; \\ \prod_{i=1}^{m'} \hat{P}^{(k-1)}(B_i), & \text{otherwise.} \end{cases} \tag{3.7}$$

As before, $n(B)$ is the number of positions in B, and $B_1, \ldots, B_{m'}$ are the subcomponents of B on scale $k-1$. It should be clear that this definition eliminates any factor

corresponding to an event where x, or a block containing x, is resonant on any scale. As explained above, those events hold by construction in the energy-following procedure, so no smallness in probability is available. We also need the corresponding weighted sum of probability bounds, *c.f.* (2.21):

$$\hat{Q}_{x,y}^{(k)} \equiv \sum_{\substack{B \text{ containing } x \text{ and } y \\ B_{x,k} \setminus B_{x,k-1} \neq \varnothing}} \hat{P}^{(k)}(B) \varepsilon^{-q_k n(B)/2}. \tag{3.8}$$

Theorem 3.2. *Given $\psi = \frac{2}{3}$, let $\chi = \psi/3$. Then for ε sufficiently small, and $q_1 \equiv \frac{1}{5}$, $q_k \equiv q_{k-1}(1 - L_{k-2}^{-\psi/20})$, we have*

$$\hat{Q}_{x,y}^{(k)} \leq \varepsilon^{q_k(|x-y| \vee L_{k-2})^\chi/2}. \tag{3.9}$$

Note that no smallness of probability can be expected for a block that consists of the single site x; smallness arises only when *other* blocks resonate with an energy from the block at x. But we see that for nontrivial blocks, the probability estimate is similar to Theorem 2.4, but with a factor of $\frac{1}{2}$ in the exponents in (3.8),(3.9). Another difference is the requirement in (3.8) that $B_{x,k} \setminus B_{x,k-1} \neq \varnothing$. This is needed because we may obtain the minimum decay length L_{k-2} only if new resonant blocks have been added to $B_{x,k-1}$ in step k.

Proof of Theorem 3.2. We follow the main steps of the proof of Theorem 2.4. For the case $k = 1$, note that $B_{x,1} \setminus B_{x,0} \neq \varnothing$ implies that $n > 1$. Hence we may counterbalance the weighting by $\varepsilon^{-q_1 n/2}$ with the probability that $n - 1$ positions other than x are resonant, *i.e.* $(3\sqrt{\varepsilon})^{n-1} \leq (3\sqrt{\varepsilon})^{n/2}$ for $n > 1$. Recalling that $q_1 = \frac{1}{5}$, we find that the sum of all trees with n sites is bounded by $(2c_d \varepsilon^{3/10})^{n/2}$. Then, since $n \geq |x - y| + 1$, we obtain (3.9) for $k = 1$. The basis of this argument is that the missing factor of $3\sqrt{\varepsilon}$ can be absorbed into the bound simply by halving the exponents, as was done in (3.8),(3.9).

When we move to the inductive step, we consider a tree graph on blocks, as in Figure 1, and repeat the argument beginning with (2.33). Only one of the scale $k - 1$ blocks $B_1, \ldots, B_{m'}$ can contain x. We use the induction hypothesis (3.9) if a block contains x, and the stronger bound (2.22) from Theorem 2.4 otherwise. However, there is no guarantee that $B_{x,k-1} \setminus B_{x,k-2} \neq \varnothing$. So we insert a partition of unity according to the value of j, the first scale at which $B_{x,j} = B_{x,k-1}$. Then $B_{x,j} \setminus B_{x,j-1} \neq \varnothing$, and we have from (3.9):

$$\sum_{j<k} \hat{Q}_{x,y}^{(j)} \leq \sum_{j<k} \varepsilon^{q_j(|x-y| \vee L_{j-2})^\chi/2} \leq \varepsilon^{q_k |x-y|^\chi/2}, \tag{3.10}$$

which is valid for $x \neq y$. (The sum over j can be controlled via the decrease $q_j \to q_k$ in the exponent.) Compared with (2.41), the exponent is halved and there is no minimum decay length. Note that $B_{x,k} \setminus B_{x,k-1} \neq \varnothing$ implies that $m' > 1$, so we may make a "halving argument" as above to improve the deficient bound (3.10) on the block containing x, at the expense of the others. Specifically, we use the inequality

$$\tfrac{1}{2}|x_1 - y_1| + |x_2 - y_2| \vee L_{k-3} \geq \tfrac{1}{2}(|x_1 - y_1| \vee L_{k-3}) + \tfrac{1}{2}(|x_2 - y_2| \vee L_{k-3}). \tag{3.11}$$

Then it should be clear that we can apply the remaining arguments proving bounds such as (2.41) on the terms $m \geq 2$. We obtain an estimate $2^{-m} \varepsilon^{q_k(|x-y| \vee L_{k-2})^\chi / 2}$ on the mth term. Similar arguments will lead to the minimum decay length L_{k-2} if $m = 1$ and $m' > 1$. The condition $B_{x,k} \setminus B_{x,k-1} \neq \varnothing$ eliminates the case $m = m' = 1$. We obtain (3.9) after summing over m. $\qquad \square$

As in Corollary 2.6, we may translate this theorem into an estimate on $\hat{\mathcal{P}}_{x,y}^{(k)}$, the probability that y belongs to the block of x on scale k. Summing (3.9) over j as in (3.10), we obtain the following result:

Corollary 3.3. *Given $\psi = \frac{2}{3}$, $\chi = \psi/3$, let ε be sufficiently small. Then there is a $q > 0$ such that*

$$\hat{\mathcal{P}}_{x,y}^{(k)} \leq \varepsilon^{q|x-y|^\chi / 2}. \tag{3.12}$$

3.3 Discriminants

The following result is important for controlling the energy-following procedure, and it will also enter into the proof of our main level-spacing theorem for H. It is an analog of (2.19) that applies to discriminants, instead of determinants. It will be used to control the probability that the level spacing of a block is less than δ. Let us put

$$\Gamma(\mathbf{u}) = \text{disc}\left(\tilde{F}_\lambda^{(k)}(B) \right), \tag{3.13}$$

where $\text{disc}(M) = \prod_{i<j} (\lambda_i - \lambda_j)^2$ is the discriminant of a matrix M with eigenvalues $\lambda_1, \ldots, \lambda_N$. It is well-known that $\text{disc}(M)$ is a homogeneous polynomial of degree $N(N-1)$ in the entries M_{ij}.

Proposition 3.4. *Let B be a block of volume n. Consider the discriminant of the Schur complement matrix $\tilde{F}_\lambda^{(k)}(B)$. For any $\delta > 0$,*

$$P\left(|\Gamma(\mathbf{u})| \leq \delta^2 3^{4n^2} \right) \leq \delta^{1/(2n^2)} \cdot 4(4n)^{n+1}. \tag{3.14}$$

Furthermore,

$$P\left(\min_{i<j} |\lambda_i - \lambda_j| < \delta \right) \leq \delta^{1/(2n^2)} \cdot 4(4n)^{n+1}, \tag{3.15}$$

where $\lambda_1, \ldots, \lambda_{2n}$ are the eigenvalues of $\tilde{F}_\lambda^{(k)}(B)$.

Proof. As explained in the proof of (2.19), the entries of $\tilde{F}_\lambda^{(k)}(B)$ are equal to those of H, plus terms of size $O(\gamma^2/\varepsilon)$ that are independent of $\mathbf{u} \equiv \{u_x\}_{x \in B}$. Hence $\Gamma(\mathbf{u})$ is a polynomial of degree $\leq 4n^2$ in $\mathbf{u}$. (Here n is the number of positions in B, so that $N = 2n$.)

We apply the Brudnyi-Ganzburg inequality (2.16) to $\Gamma(\mathbf{u})$. We need to find a point in parameter space where a lower bound on $\Gamma(\mathbf{u})$ can be proven. Take $\mathbf{u}_1 = (1, 2, \ldots, n)$. Then the eigenvalues are close to $(\pm\sqrt{2}, \pm\sqrt{5}, \ldots, \pm\sqrt{1+n^2})$. The minimum gap is

$\sqrt{5} - \sqrt{2} \approx .82$, and allowing for eigenvalue movement $O(\gamma^2/\varepsilon)$, we obtain a crude lower bound

$$\Gamma(\mathbf{u}_1) \geq (4/5)^{4n^2}. \tag{3.16}$$

Then, following the arguments for (2.17)-(2.19), we put $U = [-n, n]^n$, $\kappa = 4n^2$, and

$$\omega = \{\mathbf{u} : \Gamma(\mathbf{u}) \leq \delta^2 3^{4n^2}\}. \tag{3.17}$$

Since $\sup_U |\Gamma(\mathbf{u})| \geq (4/5)^{4n^2}$ and $|U| = (2n)^n$, we obtain

$$|\omega| \leq \left[(5/4)^{4n^2} \delta^2 3^{4n^2}\right]^{1/(4n^2)} \cdot 4n(2n)^n \leq \tfrac{15}{4} \delta^{1/(2n^2)} \cdot 2(2n)^{n+1}. \tag{3.18}$$

A factor 2^n converts this into a probability, and we obtain (3.14). Now, restricting to the support [-1,1] of the random variables u_i, the spectrum is contained in $[-\frac{3}{2}, \frac{3}{2}]$, so all eigenvalue differences are bounded by 3. Hence, if $\min_{i<j} |\lambda_i - \lambda_j| < \delta$, then $\Gamma(\mathbf{u}) < \delta^2 3^{4n^2}$. Thus (3.14) implies (3.15) and the proof is complete. $\square$

Let us relate this result to our energy-following procedure, in the context of a particular choice of $E_1, \ldots, E_k$.

Definition 3.5. *A block B of volume n is said to be **autoresonant on scale k** if $n \leq L_k^{\psi/4}$, $diam(B) \leq L_k$, and*

$$\min_{i<j} |\lambda_i - \lambda_j| < \varepsilon_k \equiv \varepsilon^{L_k^{\psi}}, \tag{3.19}$$

where $\lambda_i, i = 1, \ldots, 2n$ are the eigenvalues of $\tilde{F}_{E_k}^{(k)}(B)$.

Corollary 3.6. *Given x, $E_1, \ldots, E_k$ in the energy-following procedure, the probability that any block containing x is autoresonant on scale k is bounded by $\varepsilon^{L_k^{\psi/3}/2}$, for γ sufficiently small.*

Proof. Given the restrictions on diameter and volume, the number of possible blocks B is bounded by

$$(2L_k)^{dn} \leq e^{dL_k^{\psi/4} \log 2L_k}. \tag{3.20}$$

For a given B, (3.15) shows that the probability is bounded by

$$\varepsilon_k^{1/(2n^2)} \cdot 4(4n)^{n+1} \leq \varepsilon^{L_k^{\psi/2}/2} \cdot 4(4n)^{n+1}. \tag{3.21}$$

A combinatoric factor 2^n allows us to fix n. The corollary follows by taking the product of these bounds. $\square$

3.4 Eigenfunction Correlators

We wish to prove fractional exponential decay of $\mathbb{E} \sum_\alpha |\varphi_\alpha(y)\varphi_\alpha(z)|$, which is a strong form of localization. One should be able to prove that $\sum_\alpha |\varphi_\alpha(y)\varphi_\alpha(z)|$ is exponentially small, except on a set whose probability tends to zero rapidly with $|y - z|$. One would

introduce connectedness conditions on a rapidly growing sequence of length scales as in [15], for example. Here, we focus on decay of the averaged eigenfunction correlator, which in our procedure can decay no more rapidly than do the probability estimates, *i.e.* as a fractional exponential.

Note that the eigenvalues of H are simple, with probability 1. This can be seen by noting that the discriminant is nonzero, for a particular set of well-spaced u's (of order $|\Lambda|$), because the eigenvalues remain separated after turning on parts of H that connect different positions. The discriminant is a polynomial in $\mathbf{u}$, so it cannot vanish on a set of positive measure without being identically zero.

Theorem 3.7. *Let $\chi = \frac{2}{9}$. There exists a constant $\bar{q} > 0$ such that for γ sufficiently small, the eigenfunction correlator satisfies*

$$\mathbb{E} \sum_\alpha |\varphi_\alpha(y)\varphi_\alpha(z)| \le 4\gamma^{\bar{q}|y-z|^\chi}, \tag{3.22}$$

for any rectangle Λ.

Proof. There are three parts of the proof. In the first part, we relate the sum on α in (3.22) to the energy-following procedure. We will establish that every eigenvalue λ_0 of H can be reached via the energy-following procedure for at least one set of choices for $x, E_1, E_2, \ldots$. Specifically, there is a set of choices such that $E_k = \lambda_0$ for all k sufficiently large (depending on λ_0). Initially, we will allow ourselves the freedom to pick from amongst all the sites or blocks that have spectrum close to λ_0 at each scale. Then, we will show that there is at least one *good* choice of x such that the procedure converges to λ_0 when working with the block containing x at each scale.

We proceed step by step. In the first step, consider the set of positions x such that $\pm t_x$ is within $\varepsilon/3$ of λ_0. Clearly, this set must be nonempty, because otherwise there would be no spectrum of H_0 within $\varepsilon/3$ of λ_0, and $H = H_0 + V$ with $\|V\| \le 2d\gamma$. For any of these choices of x, we can choose $E_1 = \pm t_x$ with $|E_1 - \lambda_0| \le \varepsilon/3$. We then construct $R^{(1)}$ and $F_\lambda^{(1)}$ for $|\lambda - E_1| \le \varepsilon/2$. By the fundamental lemma, $\lambda_0 \in \text{spec } F_{\lambda_0}^{(1)}$.

Now consider what happens after having made choices for $x_1, E_1, E_2, \ldots, E_k$ such that $E_j \in \text{spec } \tilde{F}_{E_j}^{(j+1)}(B)$ and $|E_j - \lambda_0| \le \varepsilon_j/3$ for $j \le k$. Note that $|E_j - \lambda_0| \le \varepsilon_j/3$ implies that $\lambda_0 \in \text{spec } F_{\lambda_0}^{(j)}$, by repeated applications of the fundamental lemma. Let us truncate $F_{\lambda_0}^{(k)} \to \oplus_B \tilde{F}_{\lambda_0}^{(k)}(B)$. The difference has norm $\le (c_d\gamma)^{L_k^\psi} \le \varepsilon_k^3$, by Corollary 2.8. Therefore, at least one block has spectrum close to λ_0, in the sense that $\text{dist}(\lambda_0, \text{spec } \tilde{F}_{\lambda_0}^{(k)}(B)) \le \varepsilon_k^3$. For any such block, we wish to choose E_{k+1} to be a solution in $[\lambda_0 - \varepsilon_{k+1}/3, \lambda_0 + \varepsilon_{k+1}/3]$ to the equation $\lambda \in \text{spec } \tilde{F}_\lambda^{(k)}(B)$. Let us order the eigenvalues of $\tilde{F}_\lambda^{(k)}(B)$ as $\lambda_1(\lambda) \le \cdots \le \lambda_{n(B)}(\lambda)$. For some p, $|\lambda_p(\lambda_0) - \lambda_0| \le \varepsilon_k^3$. By the Lipschitz continuity of $\tilde{F}_\lambda^{(k)}(B)$ in λ (Theorem 2.9) and Weyl's inequality, the interval $[\lambda_0 - 2\varepsilon_{k+1}/3, \lambda_0 + 2\varepsilon_{k+1}/3]$ maps contractively into itself under the map $\lambda \to \lambda_p(\lambda)$. By the contraction mapping principle, there is a solution E_{k+1} to the equation $E_{k+1} = \lambda_p(E_{k+1})$ that satisfies $|E_{k+1} - \lambda_0| \le 2\varepsilon_k^3 \le \varepsilon_{k+1}/3$. This may be used to construct $R^{(k+1)}$ and $F_{E_{k+1}}^{(k+1)}$, and the procedure continues.

For $L_k > \text{diam}(\Lambda)$, there can be at most one block, and $\tilde{F}^{(k)}_\lambda(B) = F^{(k)}_\lambda$. Also, as explained above, all eigenvalues are simple, so eventually there will be exactly one solution to $\lambda \in \text{spec}\, F^{(k)}_\lambda$ in $[\lambda_0 - \varepsilon_{k+1}/3, \lambda_0 + \varepsilon_{k+1}/3]$, namely λ_0 itself. Thus we have established convergence of the procedure – at least if one is allowed the freedom to pick any of the sites/blocks that have spectrum close to λ_0 at each scale. However, we would like a somewhat stronger statement: that we can reach λ_0 starting at some x and continuing through a sequence of blocks $B_{x,j}$ containing x. To see this, start at scale k such that $\tilde{F}^{(k)}_\lambda(B) = F^{(k)}_\lambda$ and $\lambda_0 \in \text{spec}\, F^{(k)}_{\lambda_0}$. As we proceed downward in scale, we can always conclude that at least one subblock is resonant with λ_0 (in the sense that $E_{j+1} \in \text{spec}\, \tilde{F}^{(j)}_{E_{j+1}}(B)$ and $|E_{j+1} - \lambda_0| < 2\varepsilon^3_{j+1}$). The argument is the same as the one given above, except that $\tilde{F}^{(j)}_{E_{j+1}}$ (rather than $F^{(j)}_{\lambda_0}$) is the matrix that is truncated to its subblocks. Then at least one subblock has $E_j \in \text{spec}\, \tilde{F}^{(j-1)}_{E_j}$ with $|E_j - E_{j+1}| \le \varepsilon^3_j$. Proceeding down to the level of individual positions, we find one or more choices of x that can be used as a starting point for the energy-following procedure, as claimed above. This completes the first part of the proof.

Let us define $N_{x,y}$ to be the number of eigenvalues of H that can be reached via the energy-following procedure starting at x, with a resonant region that includes y. In the second part of the proof, we obtain bounds on the expectation of $N_{x,y}$ – see (3.29),(3.30) below. For each k, the procedure involves a choice of eigenvalue E_k, which we will sum over so as to account for every eigenvalue that can be reached starting at x. Once E_k is selected, there is a sum over $B_{x,k}$, the block containing x (along with an associated probability – the block sum is a partition of unity in the probability space). We may define a "stopping scale" $\underline{k}$ as the first k such that all of the following conditions hold:

1. $B_{x,k} = B_{x,\bar{k}}$, where $\bar{k}$ is the smallest integer with $L_{\bar{k}} > \text{diam}(\Lambda)$. In other words, $B_{x,k}$ has reached its maximum extent.

2. $n(B_{x,k}) \le L_k^{\psi/4}$.

3. $\text{diam}(B_{x,k}) \le L_k$.

4. $B_{x,k}$ is not autoresonant on scale k.

After scale $\underline{k}$ is reached, there is no further enlargement of $B_{x,k}$. Furthermore, conditions 2-4 imply that the eigenvalues of $\tilde{F}^{(k)}_{E_k}(B_{x,k})$ obey (3.19), *i.e.* they have a minimum spacing ε_k. This means that there is only one choice for $E_{\underline{k}+1}, E_{\underline{k}+2}, \ldots$ (because for $k \ge \underline{k}$ the intervals $[E_k - \varepsilon_{k+1}/3, E_k + \varepsilon_{k+1}/3]$ can contain no more than one solution to the equation $E_{k+1} \in \text{spec}\, \tilde{F}^{(k)}_{E_{k+1}}(B_{x,k})$.) As a result, we may bound the entirety of the choices of $E_1, E_2, \ldots$ by $[2n(B_{x,\underline{k}})]^{\underline{k}}$. (In step k, there are no more than $2n(B_{x,k})$ eigenvalues of $\tilde{F}^{(k)}_{E_k}(B_{x,k})$, hence no more than $2n(B_{x,k}) \le 2n(B_{x,\underline{k}})$ solutions to $E_{k+1} \in \text{spec}\, \tilde{F}^{(k)}_{E_{k+1}}(B_{x,k})$.) For a given set of choices $E_1, E_2, \ldots$, the probability that $B_{x,\underline{k}}$ has size n and contains y is controlled by Theorem 3.2. Likewise, Proposition 3.4 controls the probability of autoresonance.

As in the proof of Corollary 2.6, we know that $\hat{P}^{(k)}(B)$ is a bound on the probability that $B_{x,k} = B$, i.e. that B is the block containing x at step k (for a given sequence of energies $E_1, E_2, ...$ for the energy-following procedure). Note that if $\underline{k}$ is the first scale where conditions 1-4 all hold, then at least one condition fails when $k = \tilde{k} \equiv \underline{k} - 1$. Thus we consider four cases, and bound

$$\mathbb{E}\, N_{x,y} \le \mathbb{E}\, N_{x,y}^{(1)} + \mathbb{E}\, N_{x,y}^{(2)} + \mathbb{E}\, N_{x,y}^{(3)} + \mathbb{E}\, N_{x,y}^{(4)}, \tag{3.23}$$

where each term represents the case where the corresponding condition breaks down for $k = \tilde{k}$. For case 1, this means that $B_{x,\underline{k}} \setminus B_{x,\underline{k}-1} \ne \varnothing$. Let us assume $x \ne y$ for the moment. Then we may estimate

$$\mathbb{E}\, N_{x,y}^{(1)} \le \sum_{\underline{k}=1}^{\infty} \sum_{\substack{B \text{ containing } x \text{ and } y \\ B_{x,\underline{k}} \setminus B_{x,\underline{k}-1} \ne \varnothing}} [2n(B)]^{\underline{k}-1} \hat{P}^{(\underline{k})}(B)$$

$$\le \sum_{\underline{k}=1}^{\infty} \sum_{\substack{B \text{ containing } x \text{ and } y \\ B_{x,\underline{k}} \setminus B_{x,\underline{k}-1} \ne \varnothing}} (\underline{k} - 1)!\, e^{2n(B)} \hat{P}^{(\underline{k})}(B)$$

$$\le \sum_{\underline{k}=1}^{\infty} \underline{k}!\, \hat{Q}_{x,y}^{(\underline{k})} \le \sum_{\underline{k}=1}^{\infty} \underline{k}!\, \varepsilon^{q(|x-y| \vee L_{\underline{k}-2})^{\chi}/2}, \tag{3.24}$$

since for γ small $e^{2n} \le \varepsilon^{-qn/2}$, c.f. (3.8). The last inequality is from Theorem 3.2 (recall that $q_k \searrow q$ as $k \to \infty$). Noting that the exponential growth of L_{k-2}^{χ} with k dominates $k \log k$ from Stirling's formula, we have that

$$\mathbb{E}\, N_{x,y}^{(1)} \le \varepsilon^{q|x-y|^{\chi}/3}. \tag{3.25}$$

For case 2, with $\tilde{k} = \underline{k} - 1$ we bound $(2n(B))^{\tilde{k}}$ by $\tilde{k}!\, e^{2n(B)}$ as above. Then since $n(B_{x,\tilde{k}}) > L_{\tilde{k}}^{\psi/4}$, we may write

$$\mathbb{E}\, N_{x,y}^{(2)} \le \sum_{\tilde{k}=1}^{\infty} \sum_{\substack{B \text{ containing } x \text{ and } y \\ n(B) > L_{\tilde{k}}^{\psi/4}}} [2n(B)]^{\tilde{k}-1} \hat{P}^{(\tilde{k})}(B)$$

$$\le \sum_{\tilde{k}=1}^{\infty} \tilde{k}!\, \varepsilon^{q L_{\tilde{k}}^{\psi/4}/3} \sum_{j \le \tilde{k}} \hat{Q}_{x,y}^{(j)} \le \varepsilon^{q|x-y|^{\chi}/3}, \tag{3.26}$$

where we have exploited the decay in n that is built into (3.8) and applied (3.10). For case 3, we have that $\mathrm{diam}(B_{x,\tilde{k}}) > L_{\tilde{k}}$, so $B_{x,\tilde{k}}$ must contain a point z with $|z - x| > L_{\tilde{k}}$ (take $z = y$ if $|x - y| > L_{\tilde{k}}$). Then (after absorbing the factor e^{2n} into $\hat{Q}$ as above)

$$\mathbb{E}\, N_{x,y}^{(3)} \le \sum_{\tilde{k}: L_{\tilde{k}} > |x-y|} \tilde{k}! \sum_{j \le \tilde{k}} \sum_{z: |z-x| > L_{\tilde{k}}} \hat{Q}_{x,z}^{(j)} + \sum_{\tilde{k}: L_{\tilde{k}} \le |x-y|} \tilde{k}! \sum_{j \le \tilde{k}} \hat{Q}_{x,y}^{(j)}$$

$$\le \sum_{\tilde{k}=1}^{\infty} \tilde{k}!\, \varepsilon^{q(|x-y| \vee L_{\tilde{k}})^{\chi}/2.5} \le \varepsilon^{q|x-y|^{\chi}/3}. \tag{3.27}$$

33

Here we have used (3.10) again, and since in either term there is a minimum decay length $L_{\tilde{k}}$, we may use some of that decay to control the sums over z and $\tilde{k}$. Finally, in case 4, $B_{x,\tilde{k}}$ is autoresonant on scale $\tilde{k}$, which implies that $|x - y| \leq L_{\tilde{k}}$ and $n \leq L_{\tilde{k}}^{\psi/4}$ (here $n = n(B_{x,\tilde{k}})$). Therefore, $(2n)^{\tilde{k}} \leq \exp\left(\tilde{k}\log(2n)\right) \leq \exp(\tilde{k}^2)$. We apply Corollary 3.6 to bound the probability, and then

$$\mathbb{E}\, N_{x,y}^{(4)} \leq \sum_{\tilde{k}:\, L_{\tilde{k}} \geq |x-y|} e^{\tilde{k}^2} \varepsilon^{L_{\tilde{k}}^{\psi/3}/2} \leq \varepsilon^{|x-y|^{\psi/3}/3} = \varepsilon^{|x-y|^\chi/3}. \tag{3.28}$$

Combining the bounds on the four cases, we obtain

$$\mathbb{E}\, N_{x,y} \leq \varepsilon^{q|x-y|^\chi/4}, \text{ if } x \neq y. \tag{3.29}$$

If $x = y$, the above bounds apply when $\tilde{k} \geq 1$. But we also need to consider the case $\tilde{k} = 0$, which means that the block containing x contains no other points. Then there are only two eigenvalues that can be reached, and so

$$\mathbb{E}\, N_{x,x} \leq 3. \tag{3.30}$$

The third part completes the proof of Theorem 3.7. We follow the proof of (3.29),(3.30), except instead of simply counting eigenvalues, we weight them by $|\varphi_\alpha(y)\varphi_\alpha(z)|$. From (2.58) we know that any eigenfunction φ of H can be written as $G_\lambda^{(k)}\varphi^{(k)}$, where $\varphi^{(k)}$ is an eigenvector of $F_\lambda^{(k)}$ with eigenvalue λ. Here we may take $k = \underline{k} \vee \bar{k}$, which ensures that E_k is resolved and that $\tilde{F}_{E_k}^{(k)}(B) = F_{E_k}^{(k)}$, so that E_k is an eigenvalue of H.

Let $N_{x,y,z}$ denote the number of eigenvalues of H that can be reached via the energy-following procedure starting at x, and whose resonant region includes y, z. Then clearly

$$\mathbb{E}\, N_{x,y,z} \leq \mathbb{E}\, N_{x,y} \text{ and } \mathbb{E}\, N_{x,y,z} \leq \mathbb{E}\, N_{x,z}. \tag{3.31}$$

Taking the geometric mean, we obtain

$$\mathbb{E}\, N_{x,y,z} \leq \left(\mathbb{E}\, N_{x,y}\right)^{1/2}\left(\mathbb{E}\, N_{x,z}\right)^{1/2} \leq 3\varepsilon^{q|x-y|^\chi/8}\varepsilon^{q|x-z|^\chi/8}. \tag{3.32}$$

where we have applied the bounds (3.29),(3.30).

Theorem 2.10 implies that

$$\sum_\alpha |\varphi_\alpha(y)\varphi_\alpha(z)| = \sum_\alpha \left|\left(G_{\lambda_\alpha}^{(k)}\varphi_\alpha^{(k)}\right)(y)\left(G_{\lambda_\alpha}^{(k)}\varphi_\alpha^{(k)}\right)(z)\right|$$

$$\leq \sum_{x,y_1,z_1} \left(\delta_{yy_1} + \left(\frac{\gamma}{\varepsilon}\right)^{|y-y_1|^\psi}\right)\left(\delta_{zz_1} + \left(\frac{\gamma}{\varepsilon}\right)^{|z-z_1|^\psi}\right) N_{x,y_1,z_1}, \tag{3.33}$$

(Recall that $G_\lambda^{(k)}$ contains an identity matrix in the block $R^{(k)}$. This leads to the Kronecker δ terms in (3.33), as we need to consider separately the cases $y \in B_{x,k}$, $y \notin B_{x,k}$, and $z \in B_{x,k}$, $z \notin B_{x,k}$.) Taking the expectation and inserting (3.32), we obtain decay from y to z through the intermediate points y_1, x, z_1. Thus

$$\mathbb{E}\, |\varphi_\alpha(y)\varphi_\alpha(z)| \leq 4\varepsilon^{q'|y-z|^\chi}. \tag{3.34}$$

34

for some $q' < q/8$ (recall that $\chi = \psi/3$: the probability decay is slower than the graphical decay). Replacing ε with γ^ϕ, we obtain the corresponding bound in (3.22). This completes the proof of Theorem 3.7. $\qquad\square$

We may prove a refined version of Theorem 3.7 in which only spectrum in an interval $[E - \delta/2, E + \delta/2]$ is considered. The resulting bound combines the density of states estimate from Theorem 3.1 with the decay estimate just completed.

Corollary 3.8. *Let $\chi = \frac{2}{9}$. There exists a constant $\bar{q} > 0$ such that for γ sufficiently small, the bound*

$$\mathbb{E} \sum_{\alpha:\, \lambda_\alpha \in [E-\delta/2, E+\delta/2]} |\varphi_\alpha(y)\varphi_\alpha(z)| \leq \gamma^{\bar{q}(|y-z| \vee L_{\hat{k}-2})^\chi}. \tag{3.35}$$

holds for any $\delta \leq \varepsilon/2$ and any rectangle Λ. Here $\hat{k}$ is defined through the inequality $\varepsilon_{\hat{k}+1} < 2\delta \leq \varepsilon_{\hat{k}}$.

Proof. Adopting a hybrid approach, we take E to be the central energy in the Schur complement up through step $\hat{k}$. As in the proof of Theorem 3.1, we find that the probability that x is in $R^{(\hat{k})}$ is $\leq \varepsilon^{qL_{\hat{k}-2}^\chi}$, c.f. (3.3). Then we initiate the energy-following procedure, picking $x \in R^{(\hat{k})}$ and the corresponding block $B_{x,\hat{k}}$ of $R^{(\hat{k})}$. We find a solution to $\lambda \in \text{spec}\, \tilde{F}_\lambda^{\hat{k}}(B_{x,\hat{k}})$ in $[E - \varepsilon_{\hat{k}}/3, E + \varepsilon_{\hat{k}}/3]$, and continue as before. Now when the four cases are considered, c.f. (3.23), there is a minimum decay length $L_{\hat{k}-2}$, and a corresponding bound

$$\mathbb{E}\, N_{x,y}^{(i)} \leq \varepsilon^{q(|x-y| \vee L_{\hat{k}-2})^\chi/3}, \text{ for } i = 1, \ldots, 4. \tag{3.36}$$

If we consider (3.24), we have $\underline{k} \geq \hat{k}$, as all blocks are resonant on scale $\hat{k}$. In (3.26), we have $\tilde{k} \geq \hat{k}$ and $\hat{k} \leq j \leq \tilde{k}$, which also guarantees a decay of the form (3.36). Likewise, in (3.27),(3.28), we have $\tilde{k} \geq \hat{k}$, and a similar estimate holds. This completes the proof. $\qquad\square$

3.5 Level Spacing

Let us now state our main result on minimum level spacing for H. Let $\{E_\alpha\}_{\alpha=1,\ldots,2|\Lambda|}$ denote the eigenvalues of H.

Theorem 3.9. *There exists a constant $b > 0$ such that for γ sufficiently small,*

$$P\left(\min_{\alpha \neq \beta} |E_\alpha - E_\beta| < \delta\right) \leq |\Lambda| \exp\left(-b|\log\gamma|^{3/4}|\log\delta|^{1/4}\right), \tag{3.37}$$

for any rectangle $|\Lambda|$ and any $0 < \delta \leq \gamma^{1/4}/4$.

This is a theorem about microscopic level statistics: we are interested in the behavior as $\delta \to 0$. We see that the probability goes to zero faster than any power of $1/|\log\delta|$.

Proof. We modify the argument in (3.24)-(3.30) in the previous proof. Let us define $N_x(\delta)$ to be the number of eigenvalues λ of H that can be reached via the energy-following procedure starting at x, and which have another eigenvalue within δ of λ. Then noting that

$$P\Big(\min_{\alpha \neq \beta} |E_\alpha - E_\beta| < \delta \Big) \leq \sum_x \mathbb{E}\, N_x(\delta), \tag{3.38}$$

we see that a bound of the form

$$\mathbb{E}\, N_x(\delta) \leq \exp\left(-b |\log \gamma|^{3/4} |\log \delta|^{1/4} \right) \tag{3.39}$$

implies the theorem.

Let us define k by the inequality

$$\varepsilon_{k+1} < 4\delta \leq \varepsilon_k. \tag{3.40}$$

For simplicity we are assuming that $\delta \leq \varepsilon/4$ so that (3.40) holds for some $k \geq 1$. This is only a new restriction for moderately small $|\Lambda|$, in any event, since the bound (3.37) is nontrivial only for $\delta \leq \exp\left(-b^{-4} |\log |\Lambda||^4 |/| \log \gamma|^3 \right)$.

As in the proof of Theorem 3.7, the energy-following procedure generates a sum of terms – *c.f.* (3.24),(3.26),(3.27),(3.28) – that can be used to provide a bound on the expected number of eigenvalues λ satisfying some condition. Before, we had the condition that the resonant region includes x, y. Here, we have the condition that another eigenvalue is within $\varepsilon_k/4$ of λ. Let us break up the sum of terms into part A (for which $\underline{k} > k$) and part B (for which $\underline{k} \leq k$). (Recall that $\underline{k}$ is the "stopping scale," *i.e.* the first scale at which conditions 1-4 all hold.)

Consider first the part A sums. Our previous bounds on terms with $\underline{k} > k$ are already small enough for our current purposes. Thus for the part A sums, we may ignore the condition that defines $N_x(\varepsilon_k/4)$, *i.e.* that another eigenvalue is within $\varepsilon_k/4$ of the eigenvalue that is reached in the energy-following procedure. Taking $y = x$ in each of (3.24),(3.26),(3.27), (3.28), and restricting to $\underline{k} > k$ or $\tilde{k} = \underline{k} - 1 \geq k$, observe that in each case a bound is proven that indicates rapid convergence in k. Noting that $\chi = \psi/3$, the bound is $\varepsilon^{q L_k^{\psi/4}/3}$ or better in each case (times a prefactor $k!$ or $\exp(k^2)$). We may absorb the prefactor with a decrease in the coefficient in the exponent, and then we obtain a bound $\varepsilon^{q L_k^{\psi/4}/4}$ for the sum of all four cases.

Consider now part B terms, with $\underline{k} \leq k$. For such terms, all of conditions 1-4 hold at step k. In particular, $B_{x,k}$ has reached its maximum extent, and $\tilde{F}_{E_k}^{(k)}(B_{x,k})$ has a minimum eigenvalue spacing $\geq \varepsilon_k$ – see Definition 3.5. As explained in the proof of Theorem 3.7, there is a unique sequence $E_{k+1}, E_{k+2}, \ldots$ that can follow E_k in the energy-following procedure. Let λ_0 be the limiting eigenvalue. Then as explained above, $|E_k - \lambda_0| \leq \varepsilon_k^3$. We have the condition that another eigenvalue λ' of H is within $\varepsilon_k/4$ of λ_0. If that is the case, then we have $\lambda_0 \in \operatorname{spec} F_{\lambda_0}^{(k)}$, $\lambda' \in \operatorname{spec} F_{\lambda'}^{(k)}$ with $|\lambda_0 - \lambda'| \leq \varepsilon_k/4$. We claim that there is some $\tilde{\lambda}'$ with $|\tilde{\lambda}' - E_k| \leq \varepsilon_k/3$ and some level k block B' such that $\tilde{\lambda}' \in \operatorname{spec} \tilde{F}_{E_k}^{(k)}(B')$. This follows from making shifts $F_{\lambda'}^{(k)} \to F_{E_k}^{(k)} \to \oplus_B \tilde{F}_{E_k}^{(k)}(B)$ and

36

applying

$$\|F_{\lambda'}^{(k)} - F_{E_k}^{(k)}\| \le c_d \frac{\gamma}{\varepsilon}|\lambda' - E_k| \le c_d \frac{\gamma}{\varepsilon}\left(\varepsilon_k/4 + \varepsilon_k^3\right),$$

$$\|F_{E_k}^{(k)} - \oplus_B \tilde{F}_{E_k}^{(k)}(B)\| \le \varepsilon_k^3, \tag{3.41}$$

which hold by Theorem 2.9 and Corollary 2.8. Now B' cannot be $B_{x,k}$, because $\tilde{F}_{E_k}^{(k)}(B_{x,k})$ has minimum level spacing ε_k, and $E_k \in \operatorname{spec} \tilde{F}_{E_k}^{(k)}(B_{x,k})$.

The probability that both $B_{x,k}$ and B' are blocks of $R^{(k)}$ is bounded by $\hat{P}^{(k)}(B_{x,k})\tilde{P}^{(k)}(B')$. (By construction, the probabilities $(3\sqrt{\varepsilon})^{n(B)}$ and $\varepsilon^{L_k^{3\psi/5}/3}$ in (2.20),(3.7) correspond to independent events that must occur if $B_{x,k}$ and B' are both blocks of $R^{(k)}$.) So we may sum $\tilde{P}^{(k)}(B')$ over all possibilities for B' for each $B_{x,k}$ that arises in the energy-following procedure. This is bounded by the sum over y of the probability that $y \in R^{(k)} \setminus B_{x,k}$. By (3.3), this is bounded by $|\Lambda|\varepsilon^{qL_{k-2}^\chi}$. The net result is that the sum of part B terms for $\mathbb{E}N_x(\varepsilon_k/4)$ is bounded by the same four sums (3.24),(3.26),(3.27),(3.28), times an overall factor of $|\Lambda|\varepsilon^{qL_{k-2}^\chi}$. Taking $x = y$, we have by (3.30) that the combined sums are bounded by 3, and this leads to an overall bound of $3|\Lambda|\varepsilon^{qL_{k-2}^\chi}$ on the sum of part B terms.

Adding the above bounds for parts A and B, we obtain

$$\mathbb{E}\,N_x(\delta) \le \varepsilon^{qL_k^{\psi/4}/4} + 3|\Lambda|\varepsilon^{qL_{k-2}^{\psi/3}}. \tag{3.42}$$

Let us make a provisional assumption that $|\Lambda| \le \varepsilon^{-qL_k^{\psi/4}/5}$. Then we have

$$\begin{aligned}
\mathbb{E}\,N_x(\delta) &\le \varepsilon^{qL_k^{\psi/4}/4} + 3\varepsilon^{q\left[L_k^{\psi/3}2^{-2\psi/3} - L_k^{\psi/4}/5\right]} \\
&\le \varepsilon^{qL_k^{\psi/4}/4} + 3\varepsilon^{qL_k^{\psi/3}/4} \\
&\le 4\varepsilon^{qL_k^{\psi/4}/4} \le \varepsilon^{qL_k^{\psi/4}/5},
\end{aligned} \tag{3.43}$$

using $2^{-2\psi/3} - \frac{1}{5} \ge \frac{1}{4}$. Summing over $x \in \Lambda$, we obtain that

$$P\left(\min_{\alpha \ne \beta}|E_\alpha - E_\beta| < \varepsilon_k/4\right) \le |\Lambda|\varepsilon^{qL_k^{\psi/4}/5}. \tag{3.44}$$

The provisional assumption $|\Lambda| \le \varepsilon^{-qL_k^{\psi/4}/5}$ is valid in our proof of (3.44), since if it is false, (3.44) is automatically true. Recalling that $\varepsilon_k = \varepsilon^{L_k^\psi}$, we may write

$$\varepsilon^{qL_k^{\psi/4}/5} = e^{-(q/5)|\log\varepsilon||\log_\varepsilon \varepsilon_k|^{1/4}} = e^{-(q/5)|\log\varepsilon|^{3/4}|\log\varepsilon_k|^{1/4}}. \tag{3.45}$$

Then since $\delta > \varepsilon_{k+1}/4 = \varepsilon_k^{2^\psi}/4$, we have a bound $|\log\delta| \le 2|\log\varepsilon_k|$ for $\varepsilon = \gamma^\phi$ small. Thus after inserting (3.45) into (3.44), we obtain the statement of Theorem 3.9. $\qquad\square$

Remark 3. The second term on the right-hand side of (3.42) is subdominant for small δ, as shown above. If we were to keep it explicit, it would lead to a term proportional to $|\Lambda|^2$, corresponding to cases where two separate blocks are resonant to each other. If

one wishes to obtain a level-spacing bound for $\gamma^{1/4}/4 < \delta \leq 1$, that term becomes the dominant one. To see how this affects the level-spacing estimate, look directly at the level-spacing condition for $\{t_x\}_{x \in \Lambda}$, and apply (2.2):

$$P\big(\min_{x \neq y} |t_x - t_y| < 2\delta\big) \leq \tfrac{1}{2}|\Lambda|(|\Lambda| - 1) \cdot 2\sqrt{2\delta} \leq \sqrt{2\delta}|\Lambda|^2. \qquad (3.46)$$

The eigenvalues of H differ from $\{\pm t_x\}_{x \in \Lambda}$ by $O(\gamma)$, by Weyl's inequality. Hence

$$P\big(\min_{\alpha \neq \beta} |E_\alpha - E_\beta| < \delta\big) \leq \sqrt{2\delta}|\Lambda|^2, \text{ for } \gamma^{1/4}/4 < \delta \leq 1. \qquad (3.47)$$

Remark 4. We see a decrease $L_k^{\psi/3} \to L_k^{\psi/4}$ in the exponent of the probability bounds when switching from density of states estimates to level spacing estimates. This arises because the discriminant is a polynomial of degree $4n^2$, whereas the determinant has degree n. With some optimization of our procedure, one should be able to obtain exponents close to $L_k^{\psi/2}$, $L_k^{\psi/3}$, respectively, with ψ close to 1. These correspond to the worst case in a tradeoff between ε^n and $\varepsilon^{L_k^\psi/n}$ or $\varepsilon^{L_k^\psi/n^2}$, respectively.

References

[1] ANDERSON, P. Absence of diffusion in certain random lattices. *Phys. Rev. 109* (1958), 1492–1505.

[2] BACH, V. A tutorial approach to the renormalization group and the smooth Feshbach map. In: *Large Coulomb Systems*, J. Dereziński and H. Siedentop, Eds., vol. 695 of *Lecture Notes in Physics*. Springer Berlin Heidelberg, 2006, pp. 1–14.

[3] BACH, V., FRÖHLICH, J., AND SIGAL, I. Renormalization group analysis of spectral problems in quantum field theory. *Adv. Math. (N. Y). 137* (1998), 205–298.

[4] BELLISSARD, J. V., HISLOP, P. D., AND STOLZ, G. Correlation estimates in the Anderson model. *J. Stat. Phys. 129* (2007), 649–662.

[5] BOURGAIN, J. An approach to Wegner's estimate using subharmonicity. *J. Stat. Phys. 134* (2009), 969–978.

[6] BRUDNYI, A. On local behavior of analytic functions. *J. Funct. Anal. 169* (1999), 481–493.

[7] CAO, Z., AND ELGART, A. The weak localization for the alloy-type Anderson model on a cubic lattice. *J. Stat. Phys. 148* (2012), 1006–1039.

[8] CHULAEVSKY, V. From fixed-energy localization analysis to dynamical localization: An elementary path. *J. Stat. Phys. 154* (2014), 1391–1429.

[9] COMBES, J.-M., GERMINET, F., AND KLEIN, A. Generalized eigenvalue-counting estimates for the Anderson model. *J. Stat. Phys. 135* (2009), 201–216.

[10] COMBES, J. M., HISLOP, P. D., KLOPP, F., AND NAKAMURA, S. The Wegner estimate and the integrated density of states for some random operators. *Proc. Math. Sci. 112* (2002), 31–53.

[11] ELGART, A., AND SCHMIDT, D. Eigenvalue counting inequalities, with applications to Schrödinger operators. arXiv:1306.3459.

[12] ELGART, A., SHAMIS, M., AND SODIN, S. Localisation for non-monotone Schrödinger operators. *J. Eur. Math. Soc. 16* (2014), 909–924.

[13] ELGART, A., TAUTENHAHN, M., AND VESELIĆ, I. Localization via fractional moments for models on Z with single-site potentials of finite support. *J. Phys. A Math. Theor. 43* (2010), 474021.

[14] ELGART, A., TAUTENHAHN, M., AND VESELIĆ, I. Anderson localization for a class of models with a sign-indefinite single-site potential via fractional moment method. *Ann. Henri Poincaré 12* (2011), 1571–1599.

[15] FRÖHLICH, J., AND SPENCER, T. Absence of diffusion in the Anderson tight binding model for large disorder or low energy. *Commun. Math. Phys. 88* (1983), 151–184.

[16] GERMINET, F., AND KLOPP, F. Enhanced Wegner and Minami estimates and eigenvalue statistics of random Anderson models at spectral edges. *Ann. Henri Poincaré 14* (2012), 1263–1285.

[17] GERMINET, F., AND KLOPP, F. Spectral statistics for random Schrödinger operators in the localized regime. *J. Eur. Math. Soc. 16* (2014), 1967–2031.

[18] GLIMM, J., AND JAFFE, A. *Quantum Physics: A Functional Integral Point of View.* Springer, 1987.

[19] GRAF, G. M., AND VAGHI, A. A remark on the estimate of a determinant by Minami. *Lett. Math. Phys. 79* (2006), 17–22.

[20] IMBRIE, J. Z. Multi-scale Jacobi method for Anderson localization. *Commun. Math. Phys. 341* (2016), 491–521.

[21] IMBRIE, J. Z. On many-body localization for quantum spin chains. arXiv:1403.7837.

[22] KILLIP, R., AND NAKANO, F. Eigenfunction statistics in the localized Anderson model. *Ann. Henri Poincaré 8* (2006), 27–36.

[23] KLEIN, A., LENOBLE, O., AND MÜLLER, P. On Mott's formula for the ac-conductivity in the Anderson model. *Ann. Math. 166* (2007), 549–577.

[24] KLEIN, A., AND MOLCHANOV, S. Simplicity of eigenvalues in the Anderson model. *J. Stat. Phys. 122* (2006), 95–99.

[25] KLOPP, F. Localization for some continuous random Schrödinger operators. *Commun. Math. Phys. 167* (1995), 553–569.

[26] KRÜGER, H. Localization for random operators with non-monotone potentials with exponentially decaying correlations. *Ann. Henri Poincaré 13* (2011), 543–598.

[27] MINAMI, N. Local fluctuation of the spectrum of a multidimensional Anderson tight binding model. *Commun. Math. Phys. 177* (1996), 709–725.

[28] NABOKO, S., NICHOLS, R., AND STOLZ, G. Simplicity of eigenvalues in Anderson-type models. *Ark. för Mat. 51*, 1 (2011), 157–183.

[29] NAKANO, F. The repulsion between localization centers in the Anderson model. *J. Stat. Phys. 123* (2006), 803–810.

[30] TAUTENHAHN, M., AND VESELIĆ, I. Minami's estimate: Beyond rank one perturbation and monotonicity. *Ann. Henri Poincaré 15* (2013), 737–754.

[31] VESELIĆ, I. Wegner estimate and the density of states of some indefinite alloy-type Schrödinger operators. *Lett. Math. Phys. 59* (2002), 199–214.

[32] VESELIĆ, I. Wegner estimate for discrete alloy-type models. *Ann. Henri Poincaré 11* (2010), 991–1005.

[33] VESELIĆ, I. Wegner estimates for sign-changing single site potentials. *Math. Phys., Anal. Geom. 13* (2010), 299–313.

[34] WEGNER, F. Bounds on the density of states in disordered systems. *Z. Phys. B 44* (1981), 9–15.